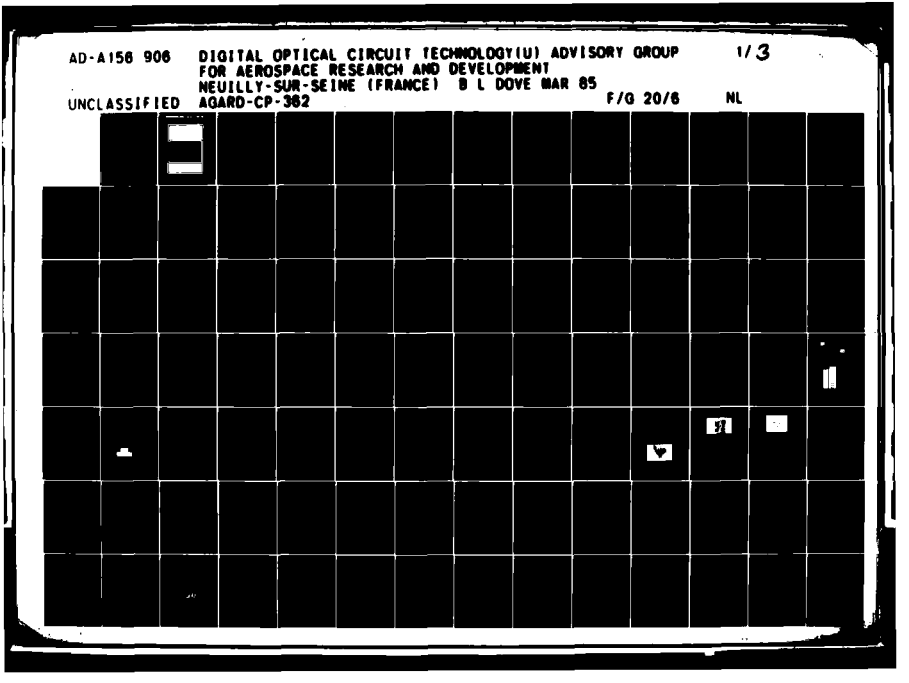
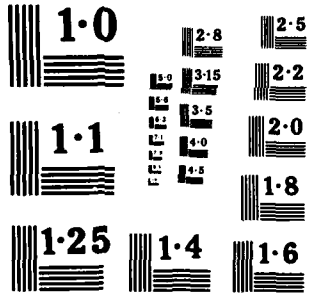
AD-A156 906 DIGITAL OPTICAL CIRCUIT TECHNOLOGY(U) ADVISORY GROUP 1/3
FOR AEROSPACE RESEARCH AND DEVELOPMENT
NEUILLY-SUR-SEINE (FRANCE) B L DOVE MAR 85
UNCLASSIFIED AGARD-CP-382 F/G 20/6 NL




1·0
1·1
1·25
2·8
3·15
3·5
4·0
4·5
1·4
2·5
2·2
2·0
1·8
1·6

AGARD-CP-362

AGARD CONFERENCE PROCEEDINGS No.362

Digital Optical Circuit Technology

DISTRIBUTION AND AVAILABILITY
ON BACK COVER

AGARD-CP-362

NORTH ATLANTIC TREATY ORGANIZATION

ADVISORY GROUP FOR AEROSPACE RESEARCH AND DEVELOPMENT

(ORGANISATION DU TRAITE DE L'ATLANTIQUE NORD)

AGARD Conference Proceedings No.362

DIGITAL OPTICAL CIRCUIT TECHNOLOGY

Edited by

Billy L.Dove
NASA Langley Research Center
Mail Stop 469
Hampton VA 23665, USA

Papers presented at the Avionics Panel Symposium held in Schliersee, Germany,
11–12 September 1984.

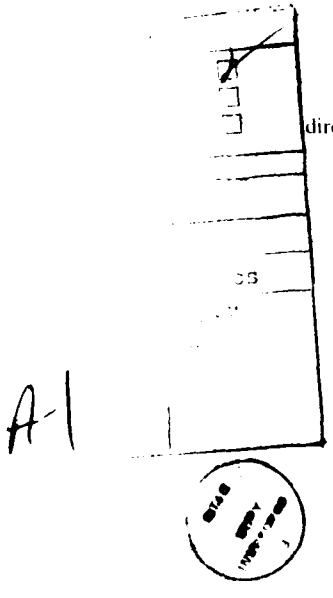
THE MISSION OF AGARD

The mission of AGARD is to bring together the leading personalities of the NATO nations in the fields of science and technology relating to aerospace for the following purposes:

- Exchanging of scientific and technical information;
- Continuously stimulating advances in the aerospace sciences relevant to strengthening the common defence posture;
- Improving the co-operation among member nations in aerospace research and development;
- Providing scientific and technical advice and assistance to the North Atlantic Military Committee in the field of aerospace research and development;
- Rendering scientific and technical assistance, as requested, to other NATO bodies and to member nations in connection with research and development problems in the aerospace field;
- Providing assistance to member nations for the purpose of increasing their scientific and technical potential;
- Recommending effective ways for the member nations to use their research and development capabilities for the common benefit of the NATO community.

The highest authority within AGARD is the National Delegates Board consisting of officially appointed senior representatives from each member nation. The mission of AGARD is carried out through the Panels which are composed of experts appointed by the National Delegates, the Consultant and Exchange Programme and the Aerospace Applications Studies Programme. The results of AGARD work are reported to the member nations and the NATO Authorities through the AGARD series of publications of which this is one.

Participation in AGARD activities is by invitation only and is normally limited to citizens of the NATO nations.



Published March 1985

ISBN 92-835-1485-8

Printed by Specialised Printing Services Limited
40 Chigwell Lane, Loughton, Essex IG10 3TZ

THEME

In the future it will be important that systems be designed which fully utilize the capabilities of optical devices. Such systems which exploit the potential speed and bandwidth capability of optical devices, and their capability for parallel processing of information should find significant applications in the communications, computing and radar fields. Technology advances of the past decade in optical electronics and very recent laboratory achievements may make possible all-optical, high-speed, EMI/EMP immune digital computers and data distribution systems. As a result of these advances in technology, there are now available sources, detectors, optical waveguides, bi-stable optical devices, modulators and demodulators capable of providing bandwidths well in excess of one gigabit. This availability is stimulating the examination of novel applications and a refinement of device performance goals is now required.

The motivation for all-optical digital systems is derived from the need to satisfy requirements for:

- Very high integrity systems
- Wide-bandwidth data distribution
- High speed
- Real-time bulk processing
- Low cost
- Elimination of optical to electronic interfaces.

The optical processing field is replete with technical approaches and to some extent there is an inadequate understanding of the organization of the problem area. This meeting on digital optical circuits will aid those on the periphery of this technology to understand its goals, approaches, and state-of-the-art. The researchers and engineers who work in this particular area had the opportunity to meet with their colleagues and discuss the technical details of this highly specialized field. The state-of-the-art realizations and potential of optical circuit technology was also of interest to a broad applications-oriented audience concerned with digital data processing, communications, radar and avionics.

Heretofore, conferences dedicated to the subject of linear or two-dimensional optical processing have overshadowed the subject of non-linear, guided wave, optical bi-stability circuits. Recent advances with non-linear and bi-stability effects have focussed attention on the possibilities for optical digital circuits. These circuits will result from a blending of non-linear materials, integrated optics technology and picosecond techniques. It is this latter subject area that the technical Specialists' Meeting on Digital Optical Circuit Technology focused upon. The Electromagnetic Wave Propagation Panel assisted in the preparation of this Meeting.

The purpose of this meeting, therefore, was to present the research and development status of optical circuit technology and to examine its relevance in the broad context of digital processing, communication, radar, avionics and flight control systems implementation.

***

Il est d'une importance essentielle que les concepteurs des systèmes futurs tirent un parti maximal des possibilités offertes par les dispositifs optiques. En effet, des systèmes exploitant le potentiel des dispositifs optiques en matière de vitesse et de largeur de bande, ainsi que leur capacité de traitement simultané de l'information, devraient trouver d'importantes applications dans les domaines des communications, du calcul et des radars. Les progrès technologiques accomplis au cours des dix dernières années au plan de l'électronique optique, et les résultats concrets tout récemment obtenus en laboratoire, permettront sans doute de réaliser des ordinateurs et des systèmes de diffusion de données qui soient à la fois numériques, entièrement optiques, très rapides et à l'abri des interférences et des impulsions électromagnétiques. Grâce à ces progrès technologiques, on dispose actuellement de sources optiques, de détecteurs, de guides d'ondes optiques, de dispositifs optiques bistables, de modulateurs et de démodulateurs capables de fournir des bandes passantes bien supérieures à un gigabit. C'es réalisations jouent un rôle stimulant dans l'étude d'applications nouvelles et il s'agit maintenant de tendre vers des performances supérieures.

La réalisation de systèmes numériques entièrement optiques répond à la nécessité de satisfaire aux impératifs suivants:

- très haut degré d'intégrité des systèmes;
- larges bandes passantes pour la diffusion des données;
- grande rapidité de fonctionnement;
- traitement de masses de données;
- coût modique;
- élimination des interfaces entre l'optique et l'électronique.

Le domaine du traitement optique est caractérisé par une multitude d'approches techniques et, dans une certaine mesure l'organisation du secteur à problèmes n'est pas parfaitement comprise. La réunion sur les circuits optiques

numériques avait pour but d'aider tous ceux qui n'ont de cette technologie qu'une connaissance encore superficielle, à comprendre ses objectifs, ses modes d'approche et son état d'avancement. Les chercheurs et ingénieurs qui travaillent dans ce secteur particulier avaient ainsi l'occasion de rencontrer leurs collègues et d'examiner avec eux les détails techniques de ce domaine hautement spécialisé. L'état de l'art en ce qui concerne les réalisations et le potentiel de la technologie des circuits optique présentait également un intérêt certain pour un auditoire orienté vers des possibilités d'applications étendues et se préoccupant de traitement numérique de données, de communications, de radar et d'avionique.

Jusqu'ici, les conférences consacrées au traitement optique linéaire ou bidimensionnel ont laissé dans l'ombre les circuits optiques bistables non linéaires à ondes guidées. Les progrès récemment accomplis en matière d'effets non linéaires et bistables ont appelé l'attention sur les possibilités des circuits. Ces circuits pourront être réalisés si l'on associe les matériaux non linéaires, la technologie de l'optique intégrée et les techniques de pico-secondes. C'est de ce domaine que traitait essentiellement la Réunion de Spécialistes de l'AVP sur la Technologie des Circuits Optiques Numériques.

Cette Réunion avait donc pour but de dresser un bilan des recherches et des réalisations intéressant la technologie des circuits optiques et d'étudier leurs implications dans le contexte général du traitement numérique des données, des communications, du radar, de l'avionique et de la mise en oeuvre des systèmes de pilotage.

AVIONICS PANEL OFFICERS

Chairman: Dr F.I.Diamond
Chief Scientist, RADC/CA
Rome Air Development Center
Griffiss AFB
N.Y. 13441
US

Deputy Chairman: Dr G.H.Hunt
ADXR (E)
Royal Aircraft Establishment
Farnborough, Hants
GU14 6TD
UK

TECHNICAL PROGRAMME COMMITTEE

Mr B.L.Dove (Chairman)
Assistant Chief
Flight Control Systems Div.
NASA Langley Research Center
Mail Stop 469
Hampton, VA 23665
USA

Prof. Ir. D.Bosman, Ne
Dr R.Klemm, Ge
Mr I.W.Mackintosh, UK
Dr E.Spitz, Fr

HOST NATION COORDINATOR

Dipl.-Ing. W.Kuny
MBB GmbH, Dept. FE 4 LKE 4
Postfach 80-11-60
8000 München 80
Germany

PANEL EXECUTIVE

Lt. Colonel T.B.Russell

CONTENTS

FOREWORD

POTENTIAL NATO APPLICATIONS FOR DIGITAL OPTICAL COMPUTERS

by

Billy L.Dove
NASA-Langley Research Center

The computing power of avionics systems has increased steadily, but is reaching the point where additional capability will be necessary to implement new functions. The conventional electronic computer has been able to keep up with (first priority) demands because of miniaturization and associated cost decreases. However, the trend in avionics is to go all-digital in the solution to complex problems, such as:

- Guidance and control systems — realized in state-space techniques requiring the manipulation of fairly large matrices.
- Variable geometry and CCV aircraft — requiring more sensors.
- Imaging and A/I techniques — increases the number of parallel data channels.
- The changing role of the man in the vehicle, from pilot to in-situ mission manager — demands more information preprocessing.

The classical von Neuman computer structure can only cope with such requirements by increasing processor speed and multiprocessor techniques. The latter, however, involve overhead in bus management so that the law of diminishing returns limits the usefulness by increasing the number of cooperating processors. The avionics community thus is in need of real parallel processor arrays, working simultaneously on the same data and on concurrent sensor signals. Several electronic solutions are being considered, but the available technology for integrating basic circuits on suitable chips almost forbids their realization in avionics formats and weights. The optical bistable devices discussed at the symposium do have the *potential* for realizing a large multichannel computer in a small package. Then not only the size, but also the power consumption of such digital optical computers can be quite modest. In some ways it even appears to be an overkill to the avionics engineering community, since 10^4 parallel channels operating at rates of gigabauds are easily within reach, and 10^6 parallel channels seem possible. In a lively discussion it was pointed out that the majority of current avionic needs could be met with 10^4 to 10^6 parallel channels.

The advice voiced at the meeting, that researchers working on optical computing should strive to join in augmenting electronic computing capability and not to compete with it in order to supplant it, should be evaluated. The infrastructure of conventional computing in both hardware and software is enormous, but deficient in capability in satisfying the most extreme computational demands. Both in support to research and design, and in the achievement of performance goals of service equipment, the demand for computational power continues to exceed capability. Examples of the former lie with aerodynamic design, and examples of the latter lie with image processing and understanding, and with artificial intelligence. Both digital optical VLSI interconnect and digital optical computing could contribute to the capability being sought. A staged approach is required, and the demonstration of digital optical computing with computationally enormous, but mathematically well-defined, problems of large matrix manipulation seems a good first step and is exemplified in the approach reported in a number of papers.

The optical digital computer is of particular interest for future applications wherever arrays of data have to be processed in parallel. It seems that the first applications of parallel computers will be devoted to special tasks such as matrix operations as used in all kinds of signal, image, and data processing. Target application areas need to be further refined. Besides the technological difficulties, an adequate use of a highly parallel optical computer could cause problems on the software side. A new "parallel" way of thinking will be necessary.

Further progress is to be expected in optical computing with the identification of useful architectures, and in so doing to focus the efforts of device innovators onto the key aspects where innovation will have more impact.

The following possible applications of the digital optical computer indicate that it is highly desirable for NATO applications. This listing, by far not complete, may give an impression of how important parallel computers are:

Signal processing — The majority of digital signal processing uses matrix algebra. In coherent systems, such as radar and sonar, complex matrix algebra is required. A fast processor operating on arrays of data that arrive in parallel (typical data rate of a radar array antenna: several thousand complex numbers per microsecond) is required to solve a number of problems, e.g., beam forming, Doppler filtering, multichannel matched filtering, clutter and jammer suppression, clutter and target recognition, synthetic aperture processing, etc. (existing array processors are serial machines!).

Radar data processing — Tracking algorithms (such as Kalman-Filter) are also based on matrix algebra. Simultaneous tracking of several targets requires parallel operation.

Image processing — Fast digital image correlators are of importance for real-time target classification and identification. A parallel digital optical computer would be the solution to this problem.

Communication switchboards — At a crossing of two bundles of N communication lines, N^2 interconnections are possible. Such an interconnection network can be realized by an optical switch array.

Multisensor signal processor — In a military environment (e.g., fighter aircraft), data coming from various sensors are obtained simultaneously and have to be processed in parallel and connected with each other.

TECHNICAL EVALUATION REPORT

Professor Hyatt M. Gibbs, Ph.D.
Optical Circuitry Cooperative and Optical Sciences Center
University of Arizona, Tucson, Arizona 85721 USA

PURPOSE AND THEME

The purpose of the meeting was to present the research and development status of optical circuit technology and to examine its relevance in the broad context of digital processing, communication, radar, avionics, and flight control systems implementation.

The following statement of the theme and objectives from the meeting announcement outlines the rationale for a reexamination of digital optical circuitry with an emphasis on all-optical systems employing nonlinear materials:

> "In the future it will be important that systems be designed which fully utilize the capabilities of optical devices. Such systems which exploit the potential speed and bandwidth capability of optical devices, and their capability for parallel processing of information should find significant applications in the communication, computing and radar fields. Technology advances of the past decade in optical electronics and very recent laboratory achievements may make possible all-optical, high-speed, EMI/EMP immune digital computers and data distribution systems. As a result of these advances in technology, there are now available sources, detectors, optical waveguides, bi-stable optical devices, modulators and demodulators capable of providing bandwidths well in excess of one gigabit. This availability is stimulating the examination of novel applications and a refinement of device performance goals is now required.
>
> The motivation for all-optical digital systems is derived from the need to satisfy requirements for:
>
> --Very high integrity systems
> --Wide-bandwidth data distribution
> --High speed
> --Real-time bulk processing
> --Low cost
> --Elimination of optical to electronic interfaces
>
> The optical processing field is replete with technical approaches and to some extent there is an inadequate understanding of the organization of the problem area. The proposed meeting on digital optical circuits will aid those on the periphery of this technology to understand its goals, approaches, and state of the art. The researchers and engineers who work in this particular area will have the opportunity to meet with their colleagues and discuss the technical details of this highly specialized field. The state-of-the-art realizations and potential of optical circuit technology are also of interest to a broad applications-oriented audience concerned with digital data processing, communications, radar and avionics.
>
> Heretofore, conferences dedicated to the subject of linear or two dimensional optical processing have over-shadowed the subject of non-linear, guided wave, optical bi-stability circuits. Recent advances with non-linear and bi-stability effects have focussed attention on the possibilities for optical digital circuits. These circuits will result from a blending of non-linear materials, integrated optics technology and picosecond techniques. It is this latter subject area that the proposed technical Specialists Meeting on Digital Optical Circuit Technology will focus upon. The Electromagnetic Wave Propagation Panel is assisting in the preparation of this Meeting."

EVALUATION

a. Materials

Nonlinear optical signal processing can best be done with very large nonlinearities that operate at room temperature, require very low power per logic element (microwatts), and are very fast (picosecond). No such material has been discovered or constructed so far. For most applications other than a large stationary computer, room-temperature operation is highly desirable if not absolutely essential. GaAs, in both bulk and multiple-quantum-well (MQW) structures, is the most promising candidate at present (talk 8; see Appendix A for talk titles and authors). It is also attractive in other ways: well-developed and commercially available diode lasers have enough power and the appropriate emission wavelength to drive GaAs nonlinear optical devices; GaAs can be used to fabricate lasers, detectors, and high-speed electronics, making it a natural material for integrated optical circuitry; a substantial base of knowledge and of growth and fabrication equipment already exists for GaAs.

Other semiconductor materials with large (≈ 1 to 10^{-5} cm^2/kW) optical nonlinearities are being investigated: InSb (5 μm, one-photon band filling, 77 K; talk 1); InSb (10 μm, two-photon band filling, 300 K; talk 1); CdS (0.49 μm, bound exciton, 4 K; talk 2); CuCl (0.39 μm, biexciton, 4 K; talk 6 and Ref. 1); InAs (3 μm, one-photon band filling, 77 K; Ref. 2) $Cd_xHg_{1-x}Te$ (10 μm, one-photon band filling, 77 K; talk 3); $Cd_xHg_{1-x}Te$ (10 μm, two-photon, 300 K; talk 3); ZnS (0.46 to 0.64 μm, thermal, 300 K; talk 8);

ZnSe (0.51 to 0.65 μm, thermal, 300 K; talks 8 and 9). Semiconductor resonantly enhanced interactions have short absorption depths (≈ 1 μm), permitting very short and fast decision making devices.

Pipeline data applications may use guided-wave devices with lengths from one to several millimeters; semiconductors can be used farther off resonance or in overlays in which only the evanescent field experiences the nonlinearity. The much longer interaction distances make the much smaller (10^{-9} to 10^{-8} cm^2/kW, i.e., very large compared with other known nonresonant nonlinearities) but faster nonlinearities of some organic materials attractive (Ref. 3).

Clearly the search for larger and faster nonlinearities, especially those still effective at room temperature, should continue.

b. Devices

Nonlinear optical devices for signal processing can be classified roughly into two categories: nonlinear etalons in which the light beams travel perpendicular to the plane of the etalon and nonlinear waveguide devices in which the light is guided through the nonlinear medium.

1. Etalons

Nonlinear etalons are attractive for parallel processing because up to a million beams could be focused on a single etalon, defining an independently functioning pixel for each beam. In the space between the nonlinear etalons the beams can be directed, imaged, or transformed by linear optical devices; the beams can pass through each other without interfering. That is, the power and advantages of optics can be fully utilized, and the nonlinear etalons make the logic decisions.

A thin film of nonlinear medium can exhibit optical bistability or thresholding, for example, by three different mechanisms: absorptive and/or dispersive bistability with external feedback or by increasing absorption bistability with only intrinsic feedback. The lowest powers, fastest speeds, and highest transmissions have been achieved using dispersive bistability. The most promising devices to date are GaAs (talk 8) etalons and ZnS (talk 8) and ZnSe (talks 1, 8, and 9) interference filters for room-temperature parallel processing.

If one imposes a maximum heat load of 100 W/cm^2, consistent with some electronic designs, one can extrapolate present-day one- or few-beam experiments as follows. An array of 10^6 spots with 10 mW per spot could run cw or up to 10 KHz on a 5-cm^2 ZnS interference filter assuming 25% absorption (i.e., up to 10^{10} operations per second). An array of 10^6 NOR gates with 1 pJ per gate could run at 100 MHz on a 1-cm^2 GaAs etalon (i.e., up to 10^{14} operations per second). Clearly lower powers are desirable, but these numbers are becoming reasonable enough to warrant the design of special-purpose few-pixel demonstrations in order to study other problems associated with parallel processing of multiple beams: crosstalk via diffraction, diffusion, or luminescence; heat dissipation; and uniformity of thickness; for example.

Nonlinear etalons are able to perform all of the basic logic operations (talks 1, 2, 3, 8, 9). Under certain conditions the transmission of a nonlinear etalon can become unstable; that is, for a noise-free steady input intensity, the output intensity can become time dependent. If the round-trip time t_R exceeds the medium response time t_M, the output may undergo very regular oscillations with a period of $2t_R$. As the values of t_M are reduced, faster and faster all-optical oscillators can be constructed (talk 4).

Most of the device research is directed toward the improvement of the most promising nonlinear etalons. But it is important to pursue studies of fundamental limits, such as determining the smallest number of nonlinear atoms required for bistability and achieving bistability without a cavity (talk 7).

2. Waveguides

Waveguides permit the maintenance of high light intensity over long propagation distances. Only the largest nonlinearities are useful for nonlinear etalons because diffraction limits the strong-focus length (the Rayleigh length = $\pi w_0^2/\lambda$, where w_0 is the beam waist at the focus and λ is the wavelength) to micrometer distances. In contrast, smaller and perhaps faster nonlinearities can be useful for guided-wave devices. Of course, the transit time increases with interaction length. But in pipelining applications, such as data encryption and decoding, transit time is unimportant provided the medium response is fast enabling a short time between pulses.

The development of all-optical waveguide bistable devices has lagged behind that for etalon devices. There have been a number of hybrid devices utilizing waveguide modulators. With fast detectors and amplifiers, hybrid devices can operate in tens of nanoseconds with very low optical powers. External amplifiers and power supplies can be eliminated if millisecond response time is acceptable. Fast hybrid devices might be the best way to harden sensitive detectors and/or eyes against lasers. High-speed optoelectronic switches may be useful in this context (talk 13). Particularly attractive are devices based on interferometric modulators such as the Mach-Zehnder directional coupler modulator (talk 10).

All-optical waveguide devices are in an early stage of development with the emphasis upon feasibility studies. The guide material can be nonlinear in principle, but the common materials for constructing waveguides are not always the most nonlinear. And the nonlinear materials often have losses preventing propagation distances of several millimeters. The solution being tried in several labs is to overlay the nonlinear material on top of a conventional waveguide. The evanescent wave can be adjusted to give the desired phase shift with acceptable losses in distances convenient for waveguide construction. Talk 5 discusses liquid CS_2 contacted to a waveguide prepared by silver-sodium ion-exchange in Schott glass type F5. Light of 1.06-μm from a Nd:YAG pulsed laser is prism-coupled or endfired into the nonlinear waveguide structure. The response time of the CS_2 should be 1 to 3 ps, but

OPTICALLY NONLINEAR AND BISTABLE BEHAVIOUR OF DIRECT GAP SEMICONDUCTORS *)

C. Klingshirn, K. Bohnert, H. Kalt and K. Kempf **)
Physikalisches Institut der Universität
Robert-Mayer-Straße 2-4
D-6000 Frankfurt am Main, Germany

SUMMARY

The optical properties of semiconductors i.e. the spectra of absorption and refraction are determined in the vicinity of the absorption edge by exciton states. At low light intensities, the optical properties of a sample are independent of the incident light power. This is the regime of linear optics. Under illumination with strong light fields, generally from lasers, the optical properties become intensity dependent. These socalled nonlinear optical phenomena can be detected by various experimental techniques like laser-induced grating spectroscopy or excite and probe beam techniques. From the experimental data, the physical origin of the nonlinearities may be deduced. The most important contributions are with increasing excitation a broadening of the exciton resonances, transitions to the biexciton and finally the phase transition to an electron-hole plasma.

Some of these excitation induced variations of the optical properties can be used to achieve optical bistability. Various types of absorptive and dispersive bistability have been found. These optically bistable devices promise interesting possibilities for application as digital memories or as elements for logical operations.

1. INTRODUCTION

One of the basic properties of crystalline semiconductors is their bandstructure, i.e. the relation between the eigenenergy E of an electron state and its quasimomentum $\vec{k}$. For an idealized direct gap semiconductor the bandstructure is shown in Fig. 1a. At zero temperature the valence band states are completely filled and the conduction band states are empty. The minimum energy distance between the bands gives the width E_g of the forbidden gap. In a socalled direct gap semiconductor the valence band maximum and

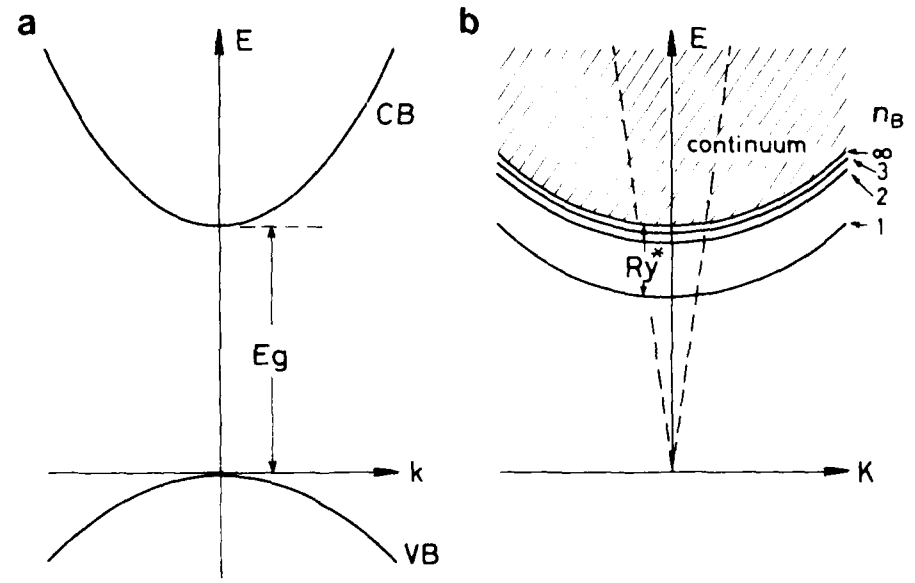


Fig. 1: Schematic presentation of the bandstructure of an idealized, direct gap semiconductor (a) and of the resulting exciton states (b). The dashed lines give the dispersion of photons in vacuum.

**) Now with IITEF, D-7800 Freiburg, Germany

*) Paper presented at the AGARD meeting on "Digital Optical Circuit Technology", Munich, September 1984

DISCUSSION

H.Gibbs, US

We obtain slightly better results with zinc sulphide than zinc selenide. Have you looked at zinc sulphide?

Author's Reply

Yes, we have indeed — and our results and yours bear some interesting comparisons because we have replaced the zinc selenide in exactly that same design with zinc sulphide — right? And we get no effect whatever. We've also done the other thing — we've replaced the spacer layer with zinc sulphide — and we do get results. That's not so surprising but it's too technical to discuss here. But there's obviously a great deal more to resolve about what's going on inside the interference filters. No doubt you'll say more later.

D.Bosman, *Ne*

How sensitive is the operation of optical processors of high bit capacity to small unintended changes of optical path lengths, e.g. due to vibration?

Author's Reply

I don't think it's very sensitive because optical devices are being made in difficult environments with laser gear for pattern recognition and all sorts of detection systems. In terms of distances and addressing things, of course optical beams hit virtually all the elements at the same time, so that's not a huge problem. So you would then be talking about vibrations over the surface area of the individual elements. Well I think you can actually make that fairly rigid. Of course it depends on the amount of resolution — the data rate you want, so that doesn't look like an impractical proposition, I think.

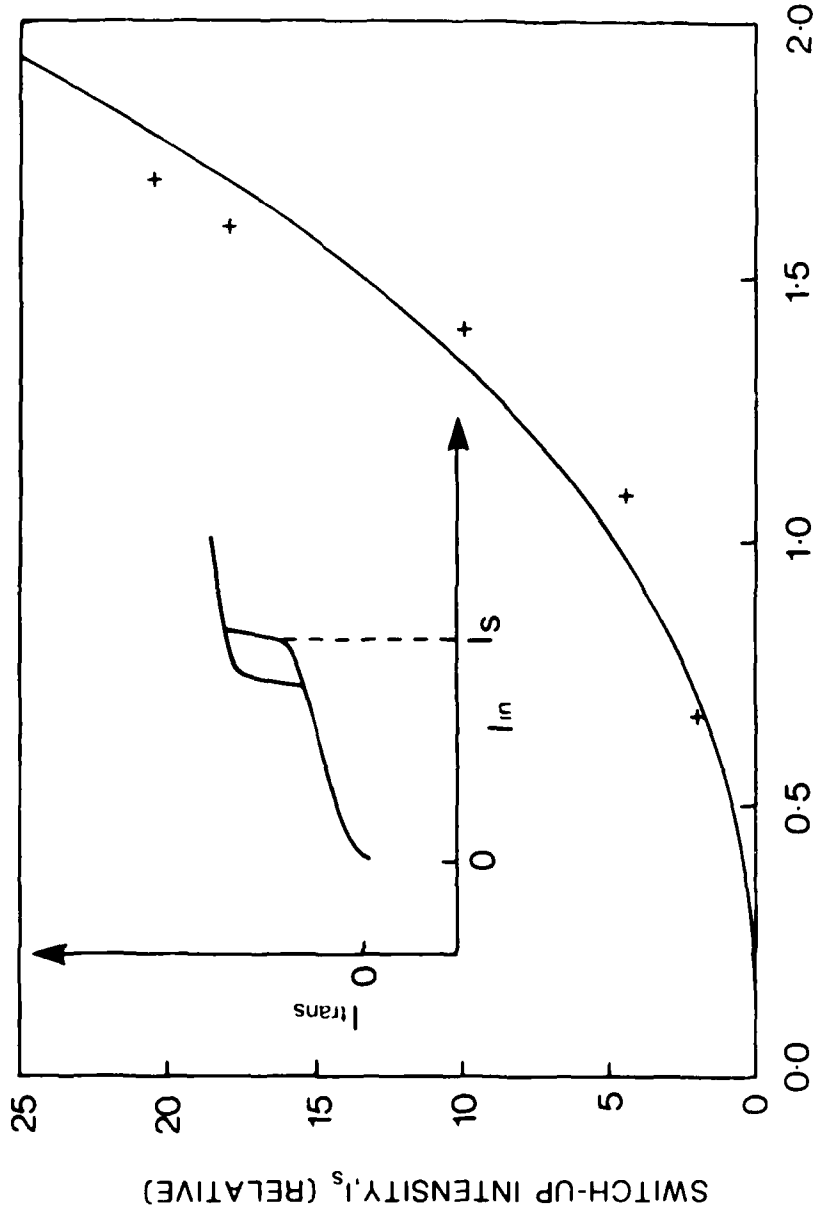


Figure 6. Switch-up irradiance as a function of initial detuning for a room-temperature InSb device operating at 10.6 μm wavelength. Inset shows a typical input/output characteristic (transmission).

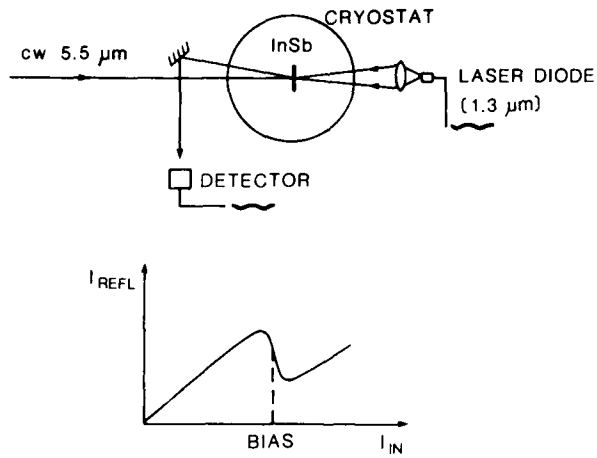


Figure 4. Experimental layout and operating condition for the demonstration of high frequency cross-wavelength modulation.

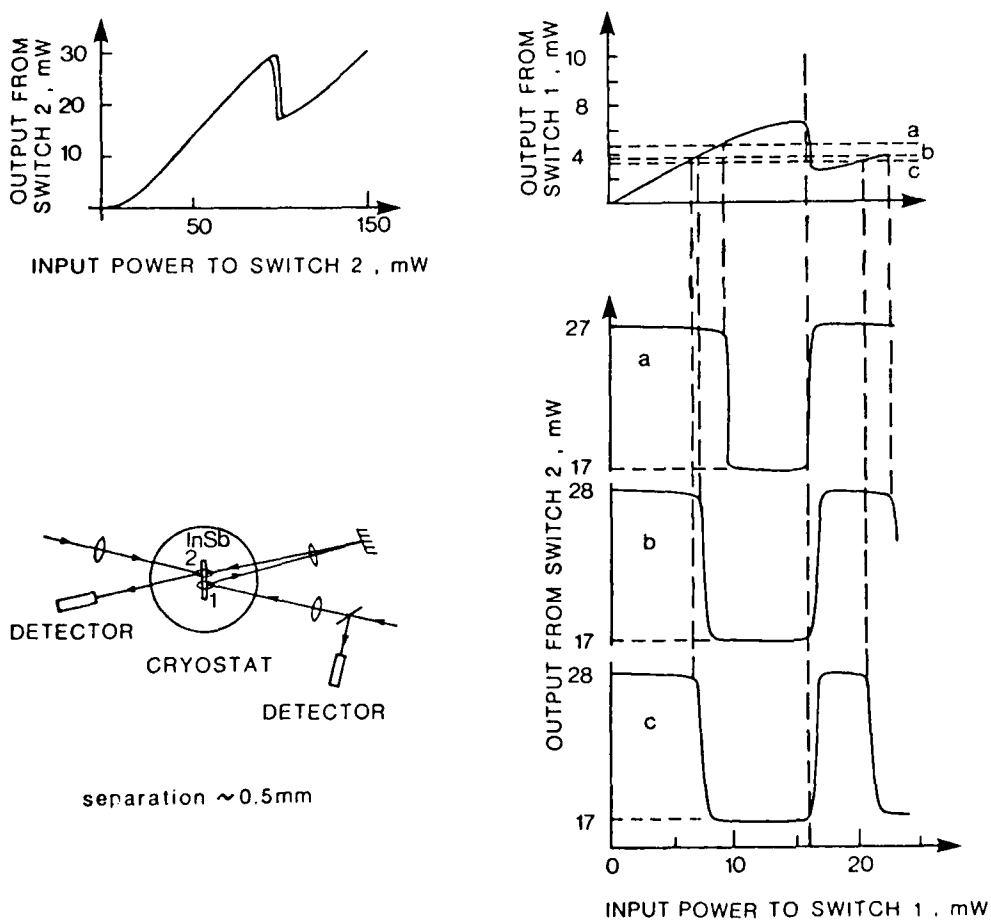


Figure 5. Two coupled OB switches on a single InSb element.
Bottom left: experimental layout.
Top: input/output characteristics of the two switches.
Bottom right: input/output characteristics of the combination.
a, b and c show the result of increasing approach of the initial bias of switch 2 to its switch point.

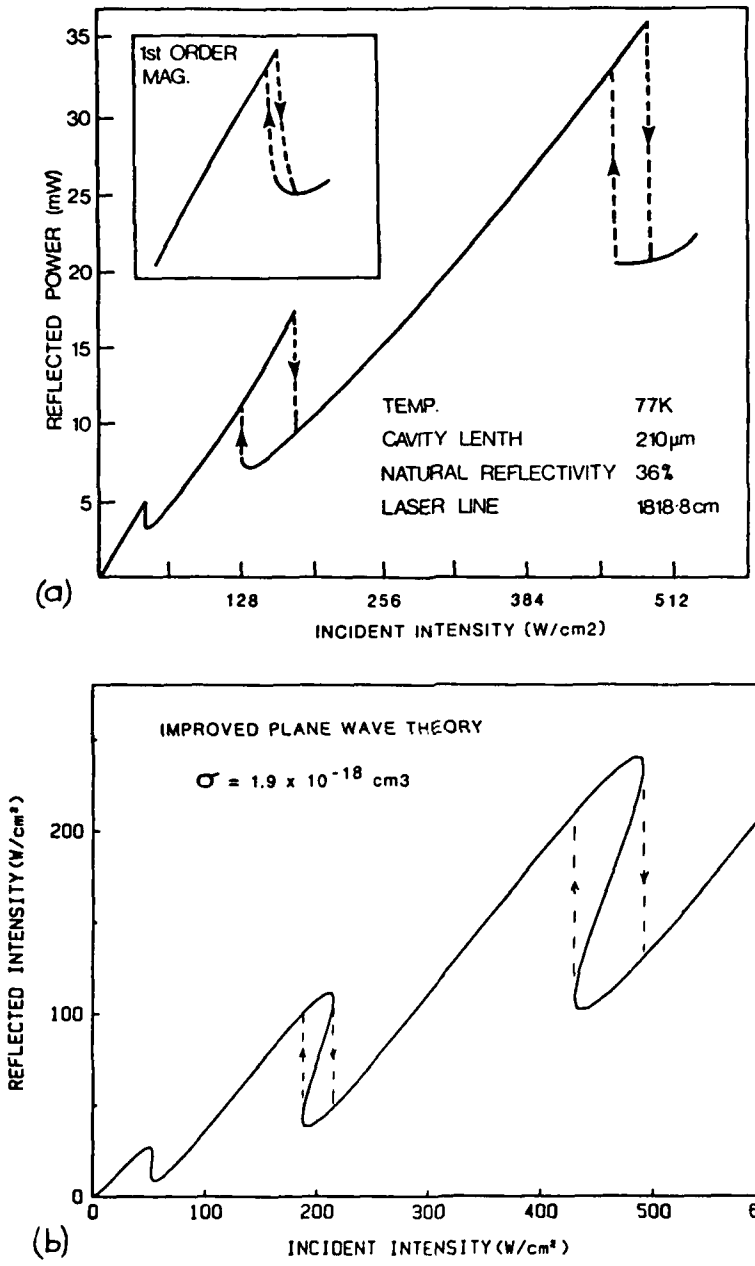


Figure 3. Input/output characteristics (reflection) for an InSb etalon. (a) Experimental result. (b) Calculated using the improved plane wave theory described in text.

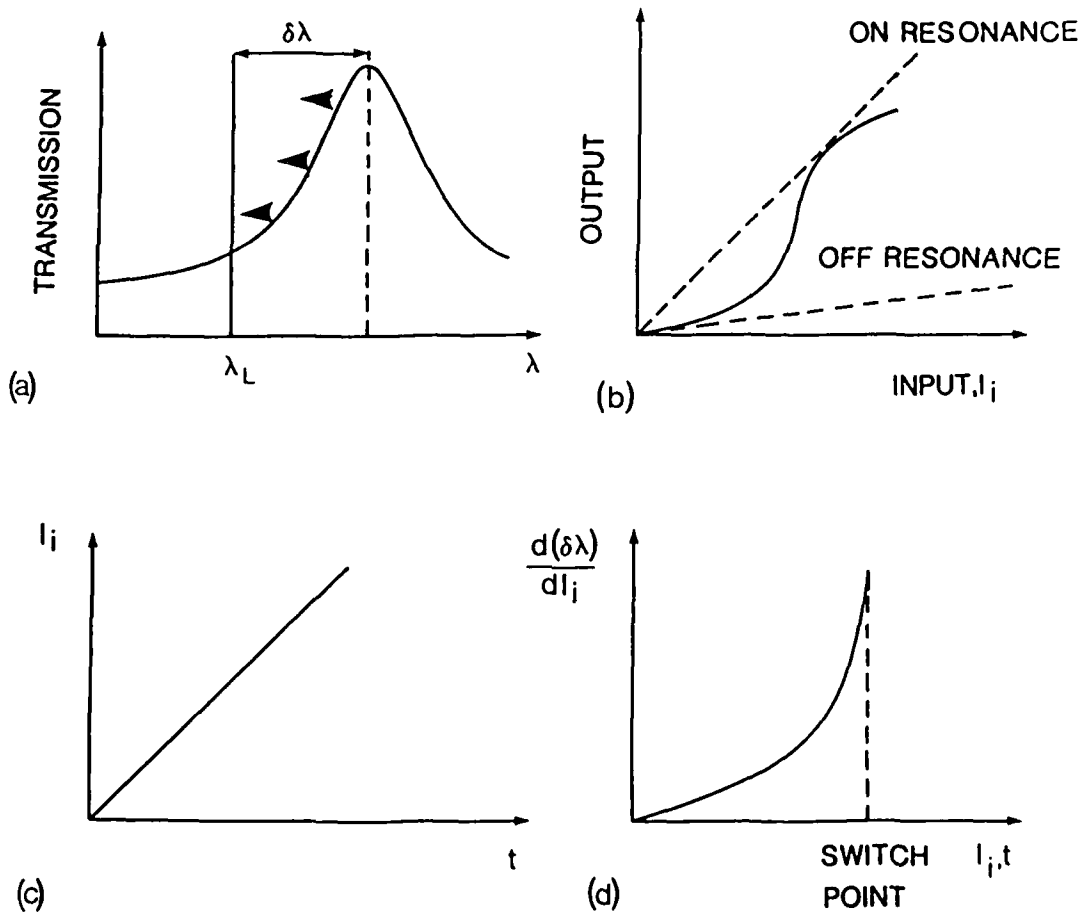


Figure 1. Showing a typical Fabry-Perot characteristic, a, and the evolution of nonlinear transmission to switching behaviour, b-d. (See text).

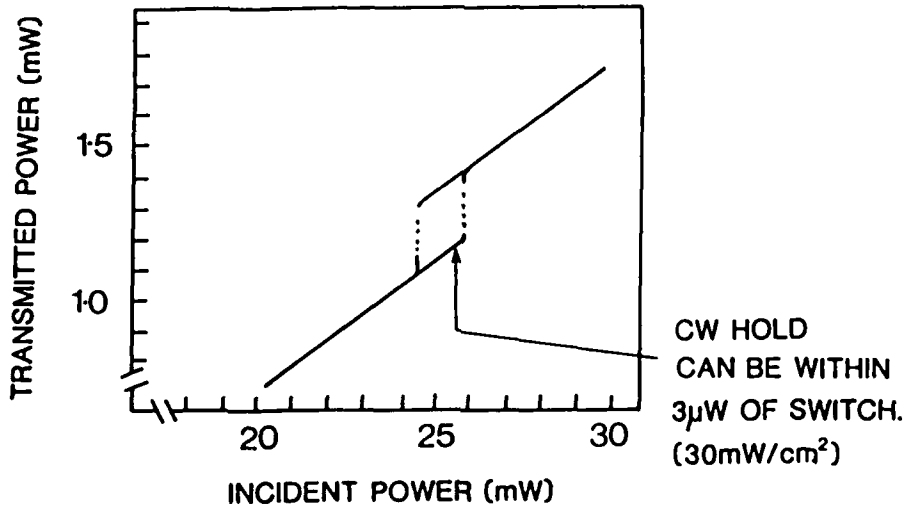


Figure 2. Input/output characteristic for a 210 μm thick InSb etalon near 77 K using 1819 cm^{-1} radiation. (Surface reflectivities 36%, spot size ~ 0.2 mm).

8. Room Temperature InSb Results

The results of initial optical bistability studies on room-temperature InSb have already been published (Kar et al. (11)). In these experiments free-carriers are excited across the 0.18 eV bandgap with 10.6 μm (CO_2 laser) radiation by a two-photon process. Recent results have been obtained using longer, ∿ 3 μs, duration pulses to minimise the dynamic effects of the rise and fall of incident power. In addition, small area pinholes mounted directly onto the samples have been used to define an area of uniform illumination at a specific point on the device. This should permit more direct comparison of experimental results with plane wave theory. Controlled variation of the initial detuning is achieved by angular adjustment of the sample. The pinhole ensures that all measurements are made on the same part of the sample, avoiding any uncertainties caused by material or surface-finish non-uniformity. Fig. 6 includes an example of the transmission characteristic which shows clear hysteresis. Care must be taken in claiming true bistability before completion of a full analysis of all the dynamic effects. Fig. 6 also plots the switching irradiance as a function of initial detuning and shows, at least qualitatively, the expected trend.

9. Conclusions

This experimental programme can be summarised under three main headings: the physics of the non-linearity and associated parameters, e.g. the form of $\Delta n(I,T)$ and $\alpha(I,T)$; the development and characterisation of single devices, e.g. switches, transphasors, modulators, etc., and the coupling of devices to develop photonic logic. Future work is directed at taking these InSb devices further including investigation of possible waveguide geometries.

It now seems possible to iterate a significant number of all-optical circuit elements and to demonstrate a simple processor.

10. References

(1) A. Szöke, V. Daneu, J. Goldhar and N.A. Kumit, Appl. Phys. Lett., 15, (1969) 376.

(2) H.M. Gibbs, S.L. McCall and T.N.C. Venkatesan, Phys. Rev. Lett., 36, (1976) 1135.

(3) E. Abraham and S.D. Smith, Reports on Progress in Physics, (8) 45, (1982) 815-885.

(4) D.A.B. Miller, M.H. Mozolowski, A. Miller and S.D. Smith, Opt. Commun., 27, (1978) 133.

(5) B.S. Wherrett, Proc. Roy. Soc. Lond. A, 390, (1983) 373.

(6) D.A.B. Miller, C.T. Seaton, M.E. Prise and S.D. Smith, Phys. Rev. Lett., 47, (1981) 197.

(7) S.W. Kurnick and J.M. Powell, Phys. Rev., 116, (1959) 597.

(8) F.A.P. Tooley, S.D. Smith and C.T. Seaton, Appl. Phys. Lett., 43, (1983) 807.

(9) C.T. Seaton, S.D. Smith, F.A.P. Tooley, M.E. Prise, and M.R. Taghizadeh, Appl. Phys. Lett., 42, (1983) 131.

(10) S.D. Smith and F.A.P. Tooley, Proc. of Topical Meeting on Optical Bistability, Rochester, U.S.A. (1983).

(11) A.K. Kar, J.G.H. Mathew, S.D. Smith, B. Davis and W. Prettl, Appl. Phys. Lett., 42, (1983) 334.

11. Acknowledgements

The authors acknowledge the assistance of Dr. H.A. MacKenzie and Mr. J. Reid in these experiments and the contribution of Mr. N. Ross to sample preparation. The 1.3 μm laser diodes were generously provided by STC Ltd., Harlow. The research was supported by both the U.K. SERC and the European Community through the EJOB Programme.

4. Transphasor Action

By adjusting the initial detuning from resonance of the InSb etalon the transmission characteristic can be made to have a steeply-sloped single-valued region - giving high differential gain. This permits the construction of a transphasor amplifier. By using a separate 3 μW, 5.5 μm wavelength chopped beam as a probe, pulse-amplification of up to 1.3×10^4 has been observed from a single device (Tooley et al. (8)). It should be noted that there is no requirement for the signal to be coherent with the bias beam in this application, as was demonstrated by using orthogonally polarised beams with equal success.

5. Response Speeds

The signal frequency in the high-gain transphasor experiment was limited to ∿ 1 KHz by mechanical chopping rates. It is important that the upper bandwidth limit of the InSb transphasor be determined. Direct observations of the switching times obtained by very slowly sweeping the input power give an upper limit of ∿ 2 μs for both switch-on and switch-off. (Much faster switch-on times should be possible with more intense, fast-rising pulses, as implied by successful switching with 35 ps Nd:YAG laser pulses (Seaton et al. (9)). To investigate the higher frequency response of a transphasor-type device a 1.3 μm laser diode, capable of being modulated at > 10 MHz has been used as the signal source. The experimental arrangement is shown in Fig. 4, together with the reflection characteristic employed. No attempt was made to achieve gain in this instance. With the device biased to the centre of the negative slope region, the signal on the 1.3 μm beam incident on the rear face of the sample induced linearly-proportional modulation on the reflected 5.5 μm beam. In initial experiments power gains of up to 20 have been observed between the 1.3 μm input and 5.5 μm output. There is evidence that the gain-bandwidth product has a constant value of about 4×10^6 Hz.

6. Cross-Wavelength Modulation and Switching

In addition to the 1.3 μm laser diode a wide range of other sources have now been used to switch or modulate the 5 μm InSb OB-switch or transphasor. These include visible wavelengths - of particular interest to potential image processing applications. For example, with the laser-diode in Fig. 4 replaced by a 10 mW He-Ne laser both switching and modulation of the 5.5 μm reflected beam by coherent visible radiation have been demonstrated. Alternatively, using a photographic flash-unit incoherent-to-coherent switching with white light has been studied (Smith et al., ibid). Table 1 summarises these and other experiments.

Table 1. Cross-wavelength modulation and switching of an InSb OB device pumped at 5.5 μm

Wavelength, etc. (Source)	Operation Mode	Gain	Frequency
5.5 μm, 3 μW (CO laser)	Switching	1.3×10^4	∿ kHz
1.3 μm, 1 mW (Laser Diode)	Switching Modulation	10	Single pulse ∿ 4 MHz
1.06 μm and 0.53 μm, 5 nJ (Nd:YAG Laser)	Switching	-	Single pulse (35 ps)
0.633 μm, 10 mW (He:Ne Laser)	Switching Modulation	∿ 1	3 kHz
White Light (Camera Flash)	Switching	-	Single pulse

7. Demonstrations of Sequentially Coupled Logic Elements

An experiment in which the transmission change through one bistable InSb device was used to switch a second has been reported previously (Smith & Tooley (10)). Recently we have coupled two devices together working in their reflection mode and, more significantly, with both gates operating adjacent to each other on a single InSb etalon. The experimental geometry is shown in Fig. 5, together with the results obtained. The two gates were addressed from opposite sides of the sample and were separated by about 0.5 mm, i.e. about 2.5 focal-spot diameters. The reflection from gate 1 was directed at gate 2. By biasing gate 2 to just below its switch point the following sequence was demonstrated, as reproduced in Fig. 4. Firstly as the input power to gate 1 is increased the reflected power becomes sufficient to switch on gate 2 (low reflection state). A further increase, however, causes gate 1 to switch on. The consequent drop in reflected power simultaneously causes gate 2 to turn off. Finally, a further increase in input power can eventually turn gate 2 back on once more. The range of input powers over which gate 2 remains on is determined by its initial bias condition. This device represents an XNOR gate and is also close to a flip-flop configuration, the latter simply requiring the output from gate 2 to be directed back at gate 1. Finally, by using transmission feedback, or by adding a third switch, an oscillator could be constructed.

Altogether, the feasibility of a wide range of logic devices have been demonstrated. For example, with a single active element only, both AND and OR gates can be made using the transmission mode, while NAND and NOR gates are obtained by operating in the reflection mode. The device described above demonstrated a two-element XNOR gate, while the first element alone acted as an XOR gate. Finally, of course, a bi-stable characteristic provides a memory element. We have now reached the point where we can start to build simple all-optical circuits.

$$I_{int} = T(\lambda) \frac{(1 + R)}{(1 - R)} I_i \quad (4)$$

so that I_{int} increases as $T(\lambda)$ rises giving positive feedback to the nonlinear shift in optical thickness. In fact, then, the peak in Fig. 1a moves increasingly quickly toward the laser wavelength λ_L. If a linear intensity ramp (with time) is used to address the device then the rate can be expressed either as a function of incident intensity or time (Fig. 1c) and is given by

$$\frac{d(\delta\lambda)}{dI_i} = T(\lambda) \Big/ \left(\frac{T_{max}}{2n_2 L} - I_i \frac{dT}{d\lambda}\right) \quad (5)$$

The two terms in the denominator of the RHS of Eq.(5) represent the nonlinear shift and feedback respectively. When the feedback, changing with I_i and $dT/d\lambda$, becomes large enough this denominator approaches zero and the rate of approach of the moving resonant wavelength to the pump wavelength diverges and becomes infinite (Fig. 1d). At this point the device switches suddenly on to resonance. The large internal field being now established, reduction in I_i leaves the resonator in its upper state and hysteretic, i.e. memory behaviour is seen. The simplest plane wave theory is obtained by simultaneous solution of Eq.(4) for $T(\lambda)$ with the standard Airy formula describing Fig. 1a,

$$T(\lambda) = \frac{1}{1 + F\sin^2\theta} \quad (6)$$

where

$$F = 4R/(1 - R)^2 \quad \text{and} \quad \theta = 2\pi nL/\lambda$$

The input-output characteristic may appear as in Fig. 1b, but varies according to choice of initial detuning $\delta\lambda$.

3. Comparison of Theory and Experiment

An example of an experimental optically bistable characteristic is given in Fig. 2 for an InSb resonator activated by a CO laser at 1819 cm^{-1}. Another characteristic, showing switching through three Fabry-Perot orders and observed in reflection, is shown in Fig. 3a. An approximate theoretical fit to these results can be achieved, without resorting to full-dimensional modelling, by including two factors in the basic plane wave theory, outlined. First, assuming that any saturation of the contribution per excess carrier-pair can be neglected,

$$\Delta n(I) = \sigma \Delta N(I) \quad (7)$$

That is, Δn may be taken as directly proportional to the density of excess carrier-pairs generated, Δn, with a constant of proportionality, σ (Miller et al. (6)). $\Delta N(I)$ then remains to be determined. This requires a knowledge of carrier recombination rates. A review of published experimental data for excess-carrier lifetimes in n-InSb at 77 - 100 K shows considerable spread of values. However, there is a clear trend showing shortening lifetimes, τ_R, at higher carrier densities ($> 5 \times 10^{15}$ cm^{-3}) and an empirical relation can be deduced of the form:

$$\tau_R^{-1} = r_1 + r_2 (N_o + \Delta N) \quad (8)$$

where N_o is the dark carrier density and ΔN the excess carrier density. This relation implies a mono-molecular, e.g. trap, recombination at low carrier densities, evolving to a bi-molecular, e.g. radiative, recombination process at higher densities. A fit to the data can be obtained using $r_1 = 1.5 \times 10^6$ s^{-1} and $r_2 = 1.5 \times 10^{-10}$ cm^3 s^{-1}, both rates being accurate to about ± 50%.

Using Eq.(8) for τ_R, the equilibrium excess carrier density can be calculated for any internal irradiance, $I(W/cm^2)$, from $\Delta N = \alpha_o \tau_R I/h\nu$. Thus $\Delta N(I)$, and hence the refractive index change, can be obtained from:

$$\Delta N(I) = \frac{(r_1 + r_2 N_o)^2 + 4r_2\alpha_o I/h\nu\,{}^{\frac{1}{2}} - (r_1 + r_2 N_o)}{2r_2} \quad (9)$$

where α_o is the carrier generating absorption coefficient (cm^{-1}) and $h\nu$ the photon energy (Joules).

It is assumed in the above that α_o is not significantly saturated at these irradiance levels. This is consistent with experimental measurements of absorption as a function of irradiance performed on other samples from the batch currently being used to fabricate bistable devices. These results demonstrate an increase in transmission losses with increasing irradiance, and appear to be consistent with a simple model based on additional absorption being induced by the generated free-carrier pairs. Assuming equal electron and hole concentrations, free-carrier absorption is dominated by the direct intra-valence-band hole transition. The hole absorption cross-section is, $\sigma_p = 2.5 \times 10^{-15}$ cm^2, at 77 K in InSb (Kurnick and Powell (7)). The total absorption coefficient is then given by:

$$\alpha = \alpha_o + \Delta N(I).\sigma_p \quad (10)$$

It has been found that by using Eq.(9) to calculate the excess carrier density, and hence both Δn from Eq.(7) and the total absorption from Eq.(10), that the input/output characteristic calculated from plane-wave theory gives a reasonable fit to the experimental result. This is shown in Fig. 3b, where σ has been taken to be 1.9×10^{-18} cm^3 (equivalent to $n_2 = 0.3$ cm^2/kW at low ΔN values). Further improvements in modelling such characteristics will probably require a fuller 2-D calculation.

OPTICALLY BISTABLE DEVICES USING InSb

S.D. Smith, F.A.P. Tooley, A.C. Walker, A.K. Kar, J.G.H. Mathew and B.S. Wherrett

Department of Physics, Heriot-Watt University, Riccarton, Edinburgh EH14 4AS, U.K.

Summary

InSb etalons operated at 77 K and illuminated by CO lasers (5.5 μm) exhibit cw optical bistability. A wide range of experiments have been performed to permit the basic characterisation of these devices and to demonstrate their various potential applications. The latter include signal amplification, modulation and, with external switching, the construction of logic gates. Two devices on a single etalon have now been coupled to form a simple all-optical circuit.

New results have also been obtained with InSb at room temperature using pulsed CO_2 lasers (10.6 μm).

1. Introduction

Abraham Szöke et al. (1) proposed that a saturable absorber inside a Fabry-Perot optical resonator could exhibit two bistable states of transmission for the same input intensity. The simple idea is that at high intensities, the induced transparency allows constructive interference at resonant wavelengths and, due to the large internal field in this situation, the system can be held 'on' to lower incident intensities than those required to induce the transparency. The experiments quoted were, however, inconclusive and it was not until 1976 that Gibbs et al. (2) demonstrated optical bistability using sodium vapour in an interferometer. They deduced that the effect was caused by nonlinear refraction rather than saturable absorption. The physics and mathematics underlying optical bistability has attracted much theoretical interest; a review was given by Abraham and Smith (3), citing 250 papers around 80% of which are theoretical.

In 1976, nonlinear refraction was explicitly observed at milliwatt powers at wavelengths near the absorption edge of the narrow gap semiconductor, InSb, in our laboratory (Miller et al. (4)). Liquid-nitrogen cooled InSb and CO lasers have proven to be an extremely useful combination in both the study of optical bistability and the demonstration of practical bistable devices. In addition, InSb with CO_2 laser illumination has permitted the demonstration of bistable switching at room temperature.

2. Theoretical Background

If we begin with intensity-dependent refraction and absorption we can simply express linear effects in intensity I by:

$$n(I) = n_o + n_2 I \tag{1}$$

where the nonlinear refractive index, n_2, can conveniently be measured in cm^2/kW, for refraction and to relate absorption coefficient $\alpha(I)$ to linear absorption α_o,

$$\alpha(I) = \alpha_o - \alpha_2 I \tag{2}$$

Both n_2 and α_2 can be described by the conventional expansion of polarisation P_i in powers of the electric field. We are concerned here with third order polarisation for intensity dependent effects proportional to $< E(\omega) >^2$ with $\omega_1 = \omega_2 = \omega_3 = \omega$. Values of $\chi^{(3)}$ reported before 1976 varied from 10^{-8} - 10^{-11} esu as reviewed by Wherrett (5). However $\chi^{(3)}$ in InSb is measured to be ~ 1 esu.

The explanation of this enormous (10^9) decrease in required power lies in the near resonance between the triply degenerate frequency (ω) of the three field components and the frequency difference between initial and various intermediate states, corresponding to the semiconductor energy gap.

The presence of absorption allows real excitation of the system: redistribution of the electron population will temporarily change the properties of the material. The nonlinearity is thus said to be 'active' and, persisting for a characteristic population lifetime τ_R (varying from microseconds to picoseconds) and can also be said to be 'dynamic'. Just as linear refraction and absorption are related by the Kramers-Kronig relation, the same integral can be used to calculate the nonlinear refraction from the change in absorption induced by population redistribution if the relaxation processes are fast compared with the measurement time. This method has been successfully used to predict the magnitude and sign of 'active' effects in semiconductors, Miller et al. (6). For band edge effects in small gap materials it is sometimes known as the 'dynamic Moss-Burstein effect' and explains both the resonance behaviour and magnitude of n_2 near (a few 10's of cm^{-1}) the band edge with values $n_2 \sim 0.1$ - 1 cm^2/kW. Since a device can be switched with an increment $\Delta n \sim 10^{-3}$, power densities of order W/cm^2 can be used in practice.

To obtain optically bistable switching feed-back is provided by a Fabry-Perot resonator, formed by the polished InSb surfaces themselves. An interferometer of optical thickness nL, shows peaks of transmission as a function of wavelength when

$$M\lambda/2 = n(I)L = (n_o + n_2 I_{int})L \tag{3}$$

where I_{int} is the intensity inside the resonator and M is integer. Consider an initial detuning from resonance of $\delta\lambda$, Fig. 1a. As the incident intensity I increases, the optical thickness changes as the internal intensity I rises, through the term $n_2 I_{int} L$ in Eq.(3). The transmission peak (Fig. 1a) thus moves towards resonance, increasing the transmission $T(\lambda)$ and thus giving a nonlinear characteristic (Fig. 1b). However the internal intensity I_{int} is related to the 'incident' intensity I by

1000-1030 BREAK

SESSION IV - ALL-OPTICAL SYSTEMS
Chairman: R. Klemm (FRG)

1030-1130 14. PROSPECTS OF THE DIGITAL OPTICAL COMPUTER
Prof. Dr. A. W. Lohmann, H. Bartelt, and J. Weigelt
Physikalishes Institut der Universitaet, Erlangen, FRG

1130-1200 15. DEFORMABLE MIRROR NEAREST NEIGHBOR OPTICAL COMPUTER
Dr. A. D. McAulay
Texas Instruments, Inc., Central Research Laboratories, Dallas, TX, US

1200-1230 16. OPTICAL TECHNIQUES FOR SIGNAL DISTRIBUTION AND CONTROL IN ADVANCED RADAR AND COMMUNICATION SYSTEMS
Prof. J. R. Forrest, Department of Electronic and Electrical Engineering, University College, London, UK

1230-1400 LUNCH

1400-1500 EEC PROJECT DISCUSSION
Prof. S. D. Smith,
Physics Department, Heriot-Watt University, Riccarton, Edinburgh, UK

1500-1530 BREAK

1530-1600 17. ELECTRO-OPTIC TECHNIQUES FOR VLSI INTERCONNECT
Dr. J. A. Neff,
Defense Advanced Research Projects Agency, DARPA/DSO, Arlington, VA, US

1600-1630 18. RESEARCH IN NUMERICAL OPTICAL COMPUTING AT THE OHIO STATE UNIVERSITY
Dr. S. A. Collins, Jr., S. F. Habiby, and A. F. Zwilling
Ohio State University, ElectroScience Laboratory, Columbus, OH, US

1630-1640 DISCUSSION OF CLOSELY COUPLED TWIN STRIPE LASERS WITH BISTABILITY
Prof. J. E. Carroll and I. H. White
University of Cambridge, UK

1640-1700 CLOSING CEREMONY

APPENDIX A

DIGITAL OPTICAL CIRCUIT TECHNOLOGY PROGRAMME

Tuesday 11 September 1984

SESSION I - OPTICAL BISTABILITY

Chairman: Prof. Ir. D. BOSMAN (NE)

0900-0930 1. OPTICALLY BISTABLE DEVICES USING InSb
Prof. S. D. Smith, F. A. P. Tooley, A. C. Walker, A. K. Kar, J. G. H. Mathew, and B. S. Wherrett,
Physics Department, Heriot-Watt University, Riccarton, Edinburgh, UK

0930-1000 2. OPTICALLY NON-LINEAR AND BISTABLE BEHAVIOUR OF DIRECT GAP SEMICONDUCTORS
Prof. Dr. C. Klingshirn, K. Bohnert, and H. Kalt,
Physikalisches Institut der Universitaet, Frankfurt, and K. Kempf, LIFEF, Freiburg, FRG

1000-1030 BREAK

1030-1100 3. OPTICAL BISTABILITY IN $Cd_xHg_{1-x}TE$
Dr. A. Miller, D. Craig, G. Parry, J. G. H. Mathew, and A. K. Kar,
Royal Signals and Radar Establishment, Malvern, Worcs, UK

1100-1130 4. OPTICAL MODULATORS AND BISTABLE DEVICES USING MOLECULAR GASES
Dr. R. G. Harrison, W. J. Firth, and I. A. Al-Saidi, and E. Cummins,
Department of Physics, Heriot-Watt University, Riccarton, Edinburg, UK

1130-1200 5. AN EXPERIMENTAL NONLINEAR OPTICAL WAVEGUIDE DEVICE
Messrs. I. Bennion, M. J. Goodwin, D. J. Robbins, and W. J. Stewart, Plessey Research (Caswell) Ltd.
Allen Clark Research Center, Towcester, Northants, UK

1200-1400 LUNCH

1400-1430 6. STATIONARY PROPERTIES AND SWITCHING CHARACTERISTICS OF DISPERSIVE OPTICAL BISTABILITY IN CuCl
Dr. C. M. Bowden, J. W. Haus, US Army Missile Laboratory, Redstone Arsenal, C. C. Sung, Department of Physics, University of Alabama, Huntsville, AL, US

1430-1500 7. CAVITYLESS OPTICAL BISTABILITY IN SYSTEMS OF TWO-LEVEL ATOMS
Dr. C. M. Bowden, U.S. Army Missile Laboratory, Redstone Arsenal, AL, and F. A. Hopf, Optical Sciences Center, University of Arizona, Tucson, AZ. US. Presented by J. W. Haus.

1500-1530 BREAK

SESSION II - OPTICAL LOGIC

Session Chairman: Mr. B. L. DOVE (US)

1530-1600 8. PROSPECTS FOR PARALLEL NONLINEAR OPTICAL SIGNAL PROCESSING USING GaAs ETALONS AND ZnS INTERFERENCE FILTERS
Prof. H. M. Gibbs, J. L. Jewell, Y. H. Lee, G. Oldbright, S. Ovadia, N. Peyghambarian, M. C. Rushford, M. Warren, and D. A. Weinberger,
University of Arizona, Tucson, AZ, US and T. Venkatesan, Bell Communications Research, Murray Hill, NJ, US

1600-1630 9. ALL-OPTICAL LOGIC GATES WITH EXTERNAL SWITCHING BY LASER AND INCOHERENT RADIATION
Dr. S. D. Smith, F. A. P. Tooley, A. C. Walker, J. G. H. Mathew, M. Taghizadeh, and B. S. Wherrett,
Physics Department, Heriot-Watt University, Riccarton, Edinburg, UK

1640-1700 10. INTEGRATED ELECTRO-OPTICAL COMPONENTS USING DIELECTRIC SUBSTRATES
Dr. M. Papuchon, Y. Bourbin, S. Vatoux and C. Puech,
Optics Department, Thomson CSF, Orsay, FR

SESSION III - SOURCES, MODULATORS, AND DEMODULATORS

Chairman: Mr. I. W. Mackintosh (UK)

0830-0900 11. THE POTENTIAL OF SEMICONDUCTORS FOR OPTICAL INTEGRATED CIRCUITS
Messers. S. Ritchie and A. G. Steventon
British Telecom Research Lab, Ipwich, UK

0900-0930 12. MULTIPORT OPTICAL DETECTORS
Mr. N. G. Walker and J. E. Caroll
Engineering Department, Cambridge University, UK

0930-1000 13. PICOSECOND PHOTOCONDUCTIVE DEVICES FOR >10 GBIT/S OPTOELECTRONIC SWITCHING
Dr. G. Veith
Standard-Elektrik Lorenz AG, Stuttgart, FRG

REFERENCES

1. N. Peyghambarian, H. M. Gibbs, M. C. Rushford, and D. A. Weinberger, "Observation of Biexcitonic Optical Bistability and Optical Limiting in CuCl," Phys. Rev. Lett. **51**, 1692 (1983). R. Levy, J. Y. Bigot, B. Hönerlage, and J. B. Grun, "Optical Bistability Due to Biexcitons in CuCl," Solid State Commun. **48**, 705 (1983).

2. [illegible]. D. Poole and E. Garmire, "Optical Bistability at the Band Gap in InAs," Appl. Phys. Lett. **44**, 363 (1984).

3. A. F. Garito and K. D. Singer, "Organic Crystals and Polymers--A New Class of Nonlinear Optical Materials," Laser Focus, p. [illegible], February, 1982.

4. H. Vach, C. T. Seaton, G. I. Stegeman, and I. C. Khoo, "Observation of Intensity-Dependent Guided Waves," Opt. Lett. **9**, 238 (1984).

5. T. Alexander, "Computing with Light at Lightning Speeds," Fortune, p. 82, July 23, 1984.

6. A. Huang, "Architectural Considerations Involved in the Design of an Optical Digital Computer," p. 780 in Ref. 7 below.

7. H. J. Caulfield, S. Horvitz, G. P. Tricoles, and W. A. von Winkle, eds. "Special Issue on Optical Computing," Proc. IEEE, July 1984.

8. A. Husain, "The Optical Interconnect: Microcircuit Problem Solver," Photonics Spectra, p. 57, August, 1984.

9. D. Casasent, "Optical Processing Research Making Significant Advancements," Laser Focus/Electro-Optics, p. 149, October 1984.

10. H. J. Caulfield, J. A. Neff, and W. T. Rhodes, "Optical Computing: the Coming Revolution in Optical Signal Processing," Laser Focus, p. 100, November 1983.

11. R. C. Alferness, "A Strong Potential for Integrated Optics Applications," Laser Focus, p. 186, October, 1984.

the response has not been studied with subnanosecond resolution. Polydiacethylene has a higher $\chi^{(3)}$ and should also be subpicosecond. Reference 4 describes earlier studies with liquid-crystal nonlinear overlays and second response times.

An attraction of using semiconductor waveguides is the potential for integrating optical sources, detectors, and electrical components on the same substrate. Talk 11 discusses the problems and possibilities in great detail and concludes that semiconductor waveguide performance should eventually be comparable to that of $LiNbO_3$, the standard material for waveguide devices.

It is anticipated that the development of nonlinear optical guided-wave devices for serial processing will be very rapid, driven by the great success of optical fibers for long-distance communications and the anticipated demand for guided-wave optical interconnects within electronic computers.

c. Optics in Computers

The progression of optics within digital computers is likely to be guided-wave interconnects, free-space imaging interconnects, special-purpose optical processors, and finally, if ever, all-optical computers.

Guided-wave interconnects can use existing technology developed for optical communication links; namely, sources, modulators, optical fibers, and detectors. Intergration of all needed components is the objective of much of the current research and development.

Much less developed or accepted is the imaging of whole arrays of information from one chip or board to another. Spatial light modulators (talk 15) and detector arrays are essential for the light-electronics interfaces. If nonlinear arrays can be used to reduce the data, then a single detector or fewer-element detector arrays would suffice. Imaging interconnects would utilize the advantages of light beams: massively parallel, no interaction between beams passing through each other, and fastest propagation speed, for example. Dynamically programmable interconnects in which beams are redirected during computation can be conceived (talk 17), emphasizing that optics offers some unique possibilities.

Special-purpose optical add-ons to electronic computers are already being marketed (Ref. 5). Some operations that optics can do better (quicker, cheaper, smaller space) are farmed out to the optics device and the answer is fed back. The evolution of electronic computers is itself in this direction of parallel architecture and dedicated subprocessors, making the resistance to the use of an optical dedicated device much lower. The study of computer architecture especially designed to fully utilize the new possibilities with optics is very important (talk 18; Ref. 6).

Perhaps someday there may even be an all-optical computer. Talk 14 discusses this possibility with considerable optimism based on several facts: the electronic computer is in trouble not because of switching speed, but because of communication problems; optics is not starting from zero, since linear operations such as imaging and Fourier transformation have been performed for many years; photons cross with no interaction except for very localized regions where "exotic" nonlinear material is carefully placed; optics is naturally suited to massive parallelism so the optical logic components can be far inferior and still yield much better systems performance.

CONCLUSIONS

It is a near certainty that optical circuit elements and devices will find commercial and military applications in an increasing number of cases. Serial data processing for communications such as telephony, data transmissions, and connections between and within computers, will surely result in greater and greater demands for high-speed integrated source/modulator/fiber-interface/detector chips. Massively parallel systems utilizing imaging and nonlinear arrays will take longer to develop but will be necessary to handle pattern recognition or guidance decisions in real time. Whether or not there will ever be an all-optical computer seems rather irrelevant. The important point is that optical circuitry certainly has contributions to make. Let's see how well it can do.

RECOMMENDATIONS FOR FUTURE EFFORTS

Research and development should be directed toward the search for better nonlinear materials, the optimization of the growth of known materials, the optimization of nonlinear devices (design, coatings, etching, or background losses may degrade performance well below that predicted for the known nonlinearity in an ideal device), improvement of light-electronics interfaces, the construction and testing of prototype systems, and the development of new algorithms and architectures taking advantage of actual and perceived optical devices. The July 1984 IEEE Proceedings Special Issue on Optical Computing (Ref. 7) is an excellent place to find many more details (see also Refs. 8-11).

Concerted efforts such as the European Joint Optical Bistability effort or the Optical Circuitry Cooperative at the University of Arizona should accelerate progress in this field.

conduction band minimum occur at the same position in k-space, generally at k = o. The bands assumed to be isotropic and parabolic can thus be described by

$$E_{CB}(k) = E_g + \frac{\hbar^2 k^2}{2m_e}$$
$$E_{vB}(k) = -\frac{\hbar^2 k^2}{2m_h} \qquad (1)$$

where m_e and m_h are the effective masses of conduction and valenceband, respectively. In an excitation process an electron is lifted from the valence into the conduction band. The empty state in the valenceband can be described as a positively charged hole with $\vec{k}_h = -\vec{k}_e$ and $\sigma_h = -\sigma_e$ where $\vec{k}_e$ and σ_e are the wavevector and the spin, respectively of the electron removed from the valenceband. The electron and hole interact via the Coulomb-potential screened by the dielectric "constant" ε. They form thus a hydrogen like series of pair states, situated below E_g. The quanta of these states are called excitons. The descrete states can be described in this approximation by

$$E(n_B,K) = E_g - Ry^* \frac{1}{n_B^2} + \frac{\hbar^2 K^2}{2M} \qquad n_B = 1,2,3 \ldots \qquad (2)$$

with: n_B : main quantum number of the hydrogen atom,

$\vec{K}$: wavevector of the exciton,

$M = m_e + m_h$: effective translational mass of the exciton,

Ry^*: reduced Rydberg energy of the exciton,

$\mu = m_e m_h (m_e + m_h)^{-1}$: reduced mass of the exciton.

The reduced Rydberg energy of the exciton is given by the one of the H-atom of 13,6 eV modified according to Eq (3)

$$Ry^* = 13{,}6 \text{ eV } \frac{\mu}{\varepsilon^2} \qquad (3)$$

In typical semiconductors Ry^* varies from 1 meV to 1oo meV. Above E_g there follow the continuum states of the exciton. The relation given in Eq (2) is visualized in Fig. 1b, together with the dispersion curve of photons (dashed line) given by

$$E = \hbar\omega = c\hbar K \qquad (4)$$

A plot of E as function of K is often used since there are conservation laws for both quantities. This E(K) plot is called a dispersion relation.

In many direct gap semiconductors the band to band transition is dipole-allowed, resulting in a strong coupling between some exciton states and photons. As a consequence a mixed state of the electronic excitation i.e. the excitons and the electromagnetic field of the photons is formed. The quanta (or quasiparticles) of this mixed state are called excitonic polaritons. The dispersion relation of these polaritons can be deduced from the dielectric function $\varepsilon(\omega,K)$ and the polariton equation. The contribution of a single excitonic resonance e.g. the $n_B = 1$ state to $\varepsilon(\omega,K)$ is given by

$$\varepsilon(\omega,K) = \varepsilon_b \left(1 + \frac{f(K)}{E(n_B = 1,K)^2 - \hbar^2\omega^2 - i\hbar^2\omega\Gamma(K)}\right) \qquad (5)$$

with ε_b : background dielectric constant, taking into account the contribution of all resonances with eigenenergies above $E(n_B = 1,K)$,

$f(K)$: oscillator strength of the resonance,

$\Gamma(K)$: damping of the resonance.

The K-dependences of f and Γ are usually neglected. The K-dependence of the eigenenergy $E(n_B = 1,K)$ is known as spatial dispersion. The polariton equation reads

$$\varepsilon(\omega,K) = \frac{c^2K^2}{\omega^2} \qquad (6)$$

Eqs (5) and (6) are an inplicite presentation of the polariton dispersion. It is shown for $\Gamma = o$ in Fig. 2a. The dispersion splits into a lower and an upper polariton branch. The longitudinal exciton branch is shown, too. The real part of the refractive index $n(\hbar\omega)$ results from Fig. 1b and 2a simply as the ratio of the polariton wavevector and the wavevector of photons in vacuum

$$n(\hbar\omega) = K_{Polariton} / k_{Photon} \qquad (7)$$

If a finite Γ is included, the evaluation of Eqs (5) and (6) yields also an imaginary part for the upper and lower polariton branches leading to a peak in the spectrum of the absorption coefficient α centered in the region of transverse and longitudinal eigenenergies.

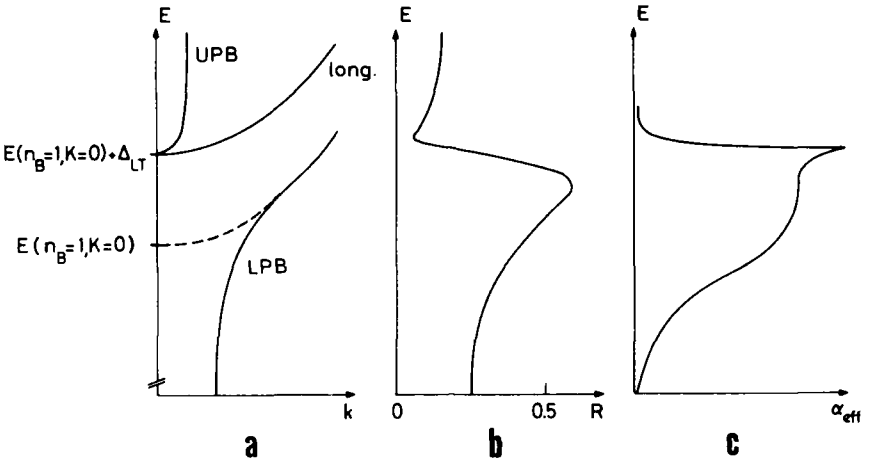


Fig. 2: The dispersion of lower and upper branch polaritons, LPB and UPB, respectively. The longitudinal exciton branch is split from the transverse eigenenergy by the longitudinal-transverse splitting Δ_{LT} (a), the reflection spectrum resulting from the dispersion (b) and the energy dependence of the effective absorption coefficient α_{eff} (c).

From the discussion above it is obvious, that the optical properties in the vicinity of the absorption edge are dominated by exciton effects. The real- and imaginary parts of the dispersion of exciton-polaritons determine the real and imaginary parts of the refractive index $\tilde{n}$

$$\tilde{n}(\omega) = n(\omega) + i\kappa(\omega) \tag{8}$$

On the other hand n and κ determine the reflection- and $\tilde{n}$ in turn is related for nonmagnetic materials to the complex dielectric function by

$$\tilde{n}^2 = \varepsilon \tag{9}$$

transmission spectra of a given semiconductor sample. A schematic picture of the polariton dispersion and the resulting reflection and absorption spectra is given in Fig. 2. The higher exciton-states (see Eq (2) for $n_B > 1$) produce similar structures, however with an oscillatorstrength decreasing as

$$f_{n_B} \sim \frac{1}{n_B^3} \tag{1o}$$

and converging towards the continuum states.

2. NONLINEAR OPTICS

The relation between an electric field $\vec{E}$ applied to matter and the resulting polarization $\vec{P}$ of the medium reads

$$\vec{P} = \varepsilon_o(\varepsilon(\omega) - 1)\vec{E} \tag{11}$$

assuming plane harmonic waves with angular frequency ω. The dependence of ε on K (Eq (5)) is neglected for simplicity. ε_o is the permeability of vacuum ($\varepsilon_o = 8.85 \cdot 1o^{12} \frac{Vs}{Am}$). As long as ε or $\tilde{n}$ are independent of $\vec{E}$ one has a linear response of the polarization $\vec{P}$ to the applied field. Correspondingly one speaks about linear optics. At sufficiently high amplitudes $\vec{E}$ or intensity I of the light field, $\varepsilon(\omega)$ itself starts to depend on $\vec{E}$. This is the field of nonlinear optics. For some recent reviews of nonlinear optics see e.g. [7-1o]. In this chapter it will be demonstrated how nonlinear optical phenomena can be detected experimentally and what are the main physical mechanisms in direct gap semiconductors leading to excitation induced variations of ε or $\tilde{n}$. The next chapter will be devoted to various possibilities by which these nonlinearities may lead to optical bistability and in the final chapter we shall consider the logic connections which can be realized by optically bistable devices. Though the phenomena, which will be discussed below are common to many direct gap semiconductors, we have selected preferentially examples for the widely known, typical direct gap semiconductor CdS, belonging to the family of II-VI compounds. In order to keep the list of references finite, we often refer to review articles, which then contain an exhaustive list of original publications. A conceptually simple and very powerful experimental technique to detect optical nonlinearities is the spectroscopy with laser induced gratings (LIG) [8,1o]. Two monochromatic, coherent and equally polarized laser-beams of generally equal frequencies are brought to spatial and temporal coincidence. In the overlap area an interference pattern is formed, i.e. there are planes of constructive interference (shown as dashed lines in Fig. 3)

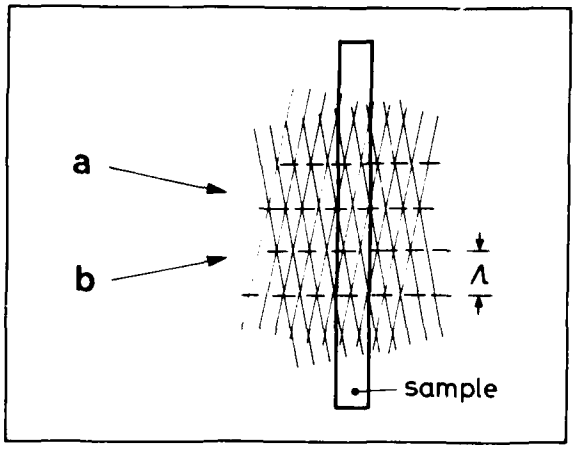


Fig. 3: Schematic drawing of the interference of two plane (laser-)light fields and the resulting grating produced in a sample

separated by regions of destructive interference. If the intensity of beam "b" I_b is much smaller than I_a the resulting intensity I will just be modulated around I_a. For $I_a = I_b$ I will vary periodically in space from $4I_a$ to o. If now a sample is brought into the interference pattern, the optical properties of which depend on the intensity and frequency of the exciting light I and $\hbar\omega_{exc}$, respectively,

$$\tilde{n}(\omega, I, \omega_{exc}) = n_o(\omega) + \Delta n(\omega, I, \omega_{exc}) + i(\kappa_o(\omega) + \Delta\kappa(\omega, I, \omega_{exc})) \tag{12}$$

a periodic modulation of $\tilde{n}$ results. An amplitude or a phase grating is formed, depending whether $\Delta\kappa$ or Δn is predominant. The grating has a spacing Λ given by the angle θ between the two beams and their wavelength λ

$$\Lambda = \frac{\lambda}{2 \sin \theta/2} \tag{13}$$

The grating decays by recombination of the excited carriers and by diffusion [7]. From this grating a further light beam may be diffracted. In the simpliest case the laser-beams which are creating the grating are diffracted themselves (see insert in Fig. 4a). The grating can be regarded as a thin (two-dimensional) one for a sample thickness d [7]

$$d << \frac{2\Lambda^2}{\lambda} \tag{14a}$$

and as a thick one, where the diffracted beams have to fulfill the Bragg condition for

$$d >> \frac{2\Lambda^2}{\lambda} \tag{14b}$$

In Fig. 4 we give the intensity of the first diffracted order in a CdS platelet of 7 µm thickness [11]. The grating is produced by the interference of two coherent dye-laser beams of about 3 nsec and o.o3 meV temporal and spatial halfwidth, respectively. They are produced by splitting the output beam of an excimer-laser pumped dye-laser into two beams of roughly equal intensity and by combining them under an angle $\theta = 1.7^o$ on the sample. This results in $\Lambda \approx 17.9$ µm fulfilling the condition for a thin grating. Diffracted orders could be observed at values as low as $I_a \approx I_b = 1oo$ W/cm^2 or $I_a = 5.1o^3$ W/cm^2 and $I_b = 1o$ W/cm^2. Easily several orders could be detected, however with rapidly decreasing intensities for increasing order indicating that the spatial variation of $\tilde{n}$ is slightly deviating from a sinusoidal one. Fig. 4a gives the spectral dependence at first order intensity 4oo W/cm^2. A LIG is formed only on the flank of the absorption edge around 2.55 eV (see also the spectrum without excitation in Fig. 6). Under these excitation conditions the damping of the exciton increases with I due to collision broadening resulting in a modification of Eq.(5) to

$$\varepsilon(\omega, K, I, \hbar\omega_{exc}) = \varepsilon_b\left(1 + \frac{f}{E(n=1,K)^2 - \hbar^2\omega^2 - i\hbar^2\omega\,\Gamma(\hbar\omega_{exc}, I)}\right) \tag{14c}$$

Thus the LIG is in this case a predominantly absorptive one. For intensities around 5 to 5o kW/cm^2 the excitation spectrum of the LIG has a different shape (Fig. 4b). It consists of a resonance-like structure with a dip in the middle. The dip is situated at 2.5497 eV, which corresponds exactly to half the biexciton energy determined independently in [12] to $E_{biex} = 5.o994$ eV. A biexciton is a quasiparticle consisting of two electrons and two holes forming a bound state with respect to two free excitons. The dip in the spectrum of Fig. 4b can thus be attributed to a two polariton transition to the biexciton, which has a high oscillatorstrength [3]. The modification of $\tilde{n}$ caused by that two polariton transition can be understood qualitatively in the following way: If light with energy $\hbar\omega_{exc}$ is sent onto a sample, polaritons of this frequency are created in the sample. In this case polaritons of energy $\hbar\omega_{abs} = E_{biex} - \hbar\omega_{exc}$ can be absorbed, converting the polaritons into

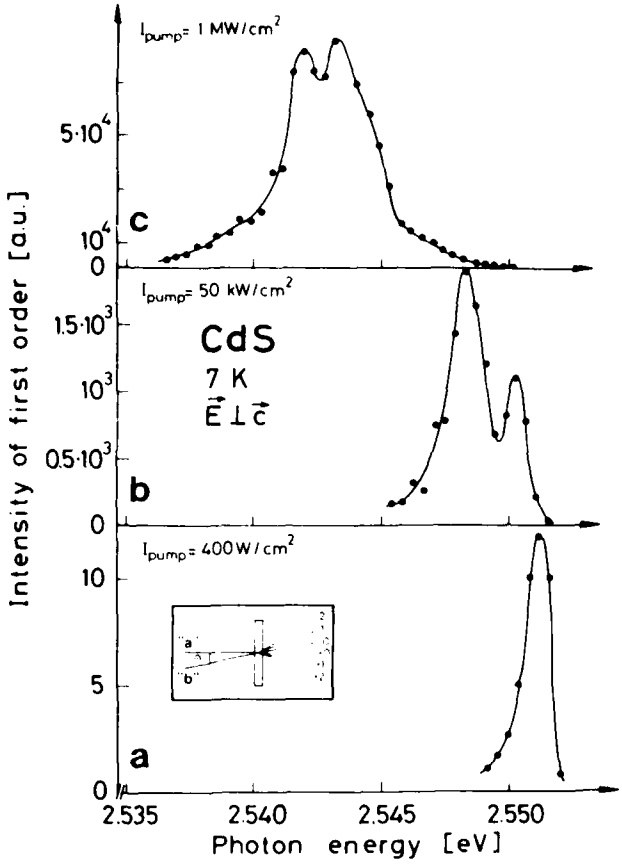


Fig. 4: The intensity of the first diffracted order + 1 from a laser induced grating in CdS as a function of the laser photon energy for three different excitation intensities. $I_{pump} = I_a + I_b$.

biexcitons. As a consequence, illumination of the sample with light of frequency $\hbar\omega_{exc}$ causes an absorption dip at $\hbar\omega_{abs}$. The oscillatorstrength of this excitation induced absorption is proportional to the density of polaritons at $\hbar\omega_{exc}$ and thus depends on I. Due to Kramers-Kronig relations, the absorption dip is connected with a resonance-like structure of the real part of the refractive index. If the dielectric function this additional resonance can be described in the simpliest approximation by

$$\varepsilon(\omega,\omega_{exc},I) = \varepsilon_b\Big(1+\frac{f}{E(n=1,K)^2-\hbar^2\omega^2-i\hbar^2\omega\Gamma(\omega,\omega_{exc},I)}+ \tag{15}$$

$$+\frac{f(I,\hbar\omega_{exc},\omega)}{(E_{biex}-\hbar\omega_{exc})^2-\hbar^2\omega^2+i\hbar^2\omega\Gamma(\omega,\omega_{exc}I)}\Big)$$

For more detailed theoretical investigations of this problem see e.g. [3,9,13] and the literature given therein. In the experiment described here, the laser is scanned over the anomaly it produces itself. The LIG is partly a dispersive, partly an absorptive one, depending on the distance $|\hbar\omega_{exc} - E_{biex}/2|$. The efficiency of the grating is roughly proportional to $I_a \cdot I_b$. In some experiments $\hbar\omega_{exc}$ and I_a has been kept constant, while I_b was varied from far $I_b \ll I_a$ to $I_b \gtrsim I_a$. Fig. 5 gives results of such experiment. I_a = 50 kW/cm^2 and $\hbar\omega_{exc}$ = 2.5505 eV have been chosen. The intensities of the beams +o, -o, +1 and -1 are plotted as a function of I_b. Beam +o is mainly the transmitted beam I_a. I_{+o} remains constant with a small decrease at the highest values of I_b due to induced absorption and diffraction. I_{-o} is the transmitted beam I_b with the admixture of the first diffracted order of I_a. I_{-o} increases linearly with I_b also with some saturation at the highest intensities. The beam I_{+1} is a superposition of the first diffracted order of I_a and the second diffracted order of I_b. Since the intensities are decreasing rapidly with increasing order, I_{+1} can be considered as mainly the first diffracted order of I_a. Since I_a is kept constant, I_{+1} gives the efficiency η of the grating which increases linearly with I_b as long as $I_b < I_a$

$$\eta \sim I_{+1}/I_a \sim \frac{d\tilde{n}}{dI}\Big|_{I_a} \cdot I_b \tag{16}$$

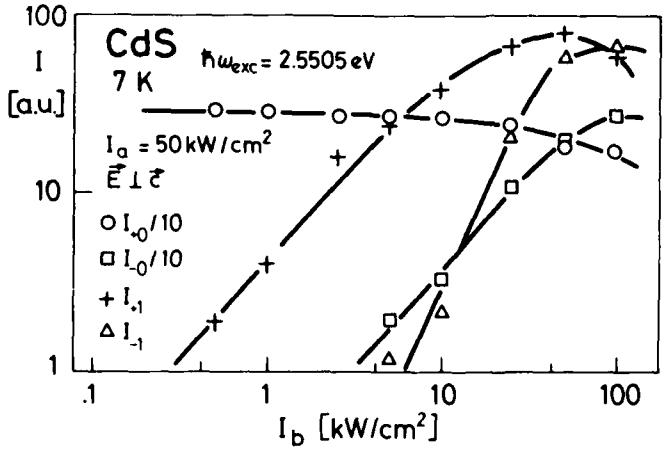


Fig. 5: As a function of the incident intensity I_b are plotted the transmitted intensities I_{+0} and I_{-0} and the diffracted intensities I_{+1} and I_{-1}. For the nomenclature see Fig. 4a.

Analoguous I_{-1} is mainly the first diffracted order of I_b resulting in Eq. (17) since I_b is varied:

$$I_{-1} \sim \eta \cdot I_b \sim I_b^2 \tag{17}$$

The experimentally observed slope of 2,1 in the logarithmis plot is in close agreement with Eq. (17). The decrease of all curves for the highest values of I_b is due to the increase of induced absorption.

At still higher excitation intensities around 1MW/cm^2 the excitation spectrum of LIG is shifted even further to the red and is rather broad and unstructured. The small dip around 2.543 eV is just due to the Fabry-Perot modestructure of the platelet type sample. At these excitation conditions, the LIG is due to the formation of an electron-hole plasma (EHP). Its properties shall be shortly outlined here: If electron-hole pairs are excited in a semiconductor at low density and temperature, they will form excitons. If the generation rate is increased, the density of excitons may become so high, that their mean distance is comparable to their Bohr-radius. In this case excitons are no longer individual quasiparticles. The exciton concept breaks down and a new collective phase is formed, the EHP. This transition from an insulating exciton phase to a metallic plasma phase is referred to as Mott transition.

This phase transition has distinct consequences for the bandstructure. It has been found both theoretically and experimentally [3,13,14-16] that the width of the forbidden gap is a monotoneously decreasing function of the electron-hole pair density in the plasma n_p. The chemical potential μ of the plasma (i.e. the energetic distance between the quasi-Fermi levels of electrons and holes) strongly depends on both n_p and the plasma temperature T_p. The analysis of the functions $E_g'(n_p)$ and $\mu(n_p,T_p)$ shows that there is a first order phase-transition from a low density exciton gas to a plasma liquid under quasiequilibrium conditions below a critical temperature T_c in close analogy to a real gas. In indirect gap semiconductors like Ge or Si this phase transition has nicely been observed [17]. In direct gap materials like CdS the EHP does not reach its liquid-like state due to the short carrier lifetime [16].

A drastic variation of the optical properties is connected with the formation of an EHP. The exciton-resonance disappears from the reflection and transmission spectra [3,16,18]. At low temperatures $\mu(T_p,n_p)$ may become larger than $E_g'(n_p)$. As a consequence, the sample is strongly absorbing due to band to band transitions above μ but between μ and E_g there is optical amplification of light i.e. gain due to the population inversion. Below E_g' the sample is simply transparent [3,13,14,16,18]. These variations cannot be simply described by adding a term to the dielectric function. They need a more sophisticated theoretical treatment, which is beyond the scope of this review. The reader is referred e.g. to [9,13] and the literature cited therein. The best experimental technique to observe the above mentioned changes of the optical spectra is the excite and probe technique. This is the second widely used possibility to detect optical nonlinearities, which will be described here: A weak, spectrally broad probebeam is sent onto the sample and the transmission- or reflection spectrum is measured once without and one with simultaneous illumination of the sample by an intense, spectrally narrow excitation pulse. The change of the reflection- and transmission spectra can be evaluated and the variations of the optical properties can be deduced. Recently this method has been improved to allow also spatially and/or temporally resolved investigations [18].
Fig. 6 gives an example, where the above mentioned variations of the optical properties connected with an EHP can be seen. The spectrum without additional excitation shows the absorptionband due to the $n_B = 1$ A_5 exciton state around 2.552 eV. The modulation of the spectrum results from the Fabry-Perot interferences in the platelet type sample. The quantitative analysis of the maxima and minima gives the real part of the refractive index $n_0(\omega)$ (see Fig. 7b), the transmission spectrum itself the absorption coefficient α connected with α by

$$\alpha = 2\kappa \frac{\omega}{c} \tag{18}$$

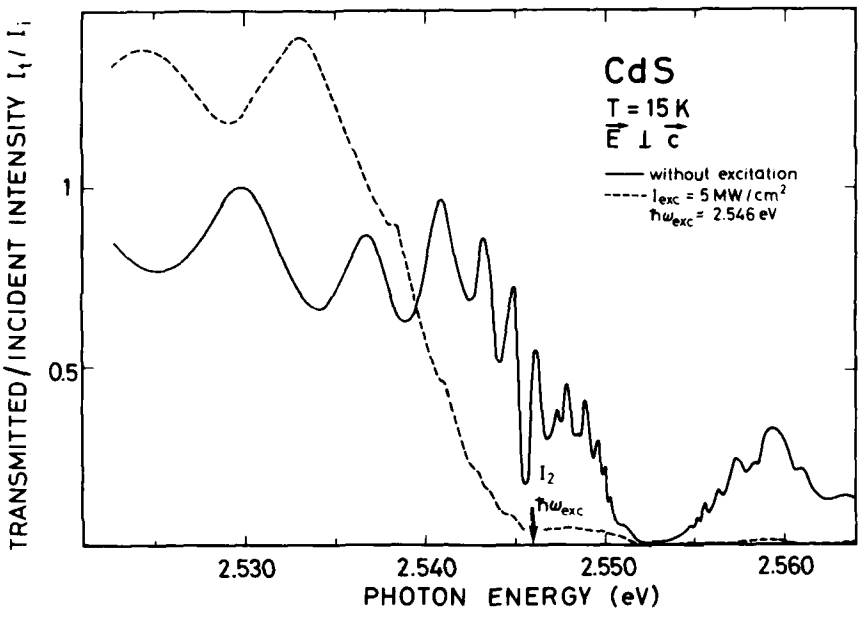


Fig. 6: Transmission spectrum of a 2.5 µm thick CdS platelet without excitation and with additional excitation.

The reflection spectrum which is not shown here exhibits the excitonic reflection-structure around 2.552 eV (see e.g. [16,18]). The second spectrum in Fig. 6 shows the transmission, while the sample is excited with a laserpulse of $I = 5$ MW/cm^2 and $\hbar\omega_{exc} = 2.546$ eV. One finds optical amplification at photon energies below 2.54 eV and the onset of absorption above, resulting in $\mu \approx 2.54$ eV. The exciton-resonance disappears both from the transmission spectrum (Fig. 6) and the reflection spectrum [3,16,18]. The evaluation of these data yields $\Delta\alpha(\hbar\omega, \hbar\omega_{exc}, I)$. The blue shift of the Fabry-Perot modes indicates a $\Delta n(\omega, I_{exc}, \hbar\omega_{exc}) < 0$ [19]. The data in Fig. 6 allow to determine the variation of n and α in the spectral region around the exciton resonance for a given $\hbar\omega_{exc}$ and I_{exc}. By extrapolating $\hbar\omega \to \hbar\omega_{exc}$ the variations at the position of the laser itself can be deduced. If $\hbar\omega_{exc}$ is varied, one gets $\Delta n(\hbar\omega = \hbar\omega_{exc}, I_{exc})$ and $\Delta\alpha(\hbar\omega = \hbar\omega_{exc}, I_{exc})$ [20]. Corresponding results are shown in Fig. 7a and c. The absorption increases below the free exciton resonance down to values of $\hbar\omega_{exc}$ around 2.54 eV. The refraction index decreases below 2.54 eV because of the disappearance of the exciton resonance, however with the onset of a resonance-like structure around 2.54 eV i.e. in the vicinity of the onset of absorption. More details of these types of experiments, which also agree with results from gain- and reflection excitation-spectroscopy [16], will be published elsewhere. A comparison of Fig. 4c and Fig. 7a,c finally reveals, that the LIG is under these excitation conditions connected with the formation of an EHP. For photon energies above 2.54 eV the grating is predominantly an absorptive one, and a dispersive one below. The main plasma-parameters deduced for CdS by various experiments are [19,20,21] for a lattice temperature around 5K:

plasma density n_p: $10^{18} \lesssim n_p \lesssim 4 \cdot 10^{18}$ cm^{-3} increasing with I_{exc}

plasma temperature T_p: $10K \lesssim T_p \lesssim 40K$

plasma lifetime τ_p: $\tau_p \approx (150 \pm 50)$ps

chemical potential μ: $\mu \approx 2.54$ eV

drift distance under inhomogeneous excitation conditions l_D: $l_D \approx (3 \pm 2)$µm.

To conclude this section, we may state, that there exist numerous processes in semiconductors which may lead to excitation induced variations of the optical properties. In the next chapter it will be shown, how these nonlinearities can be used for the realization of optical bistability.

3. OPTICAL BISTABILITY

An optically bistable device has two reversible, stable states of high and low transmission or reflection, the existence of which depends on the history. In semiconductors, there are two types of optical bistability (OB) which make use of the feedback in a Fabry-Perot cavity (FP). This cavity is usually formed from the uncoated or coated surfaces of the semiconductor sample itself. The dispersive OB depends on the variation of the real part of the refractive index $\Delta n(\omega_{exc}, I_{exc})$ of the semiconductor in the resonator. Fig. 8

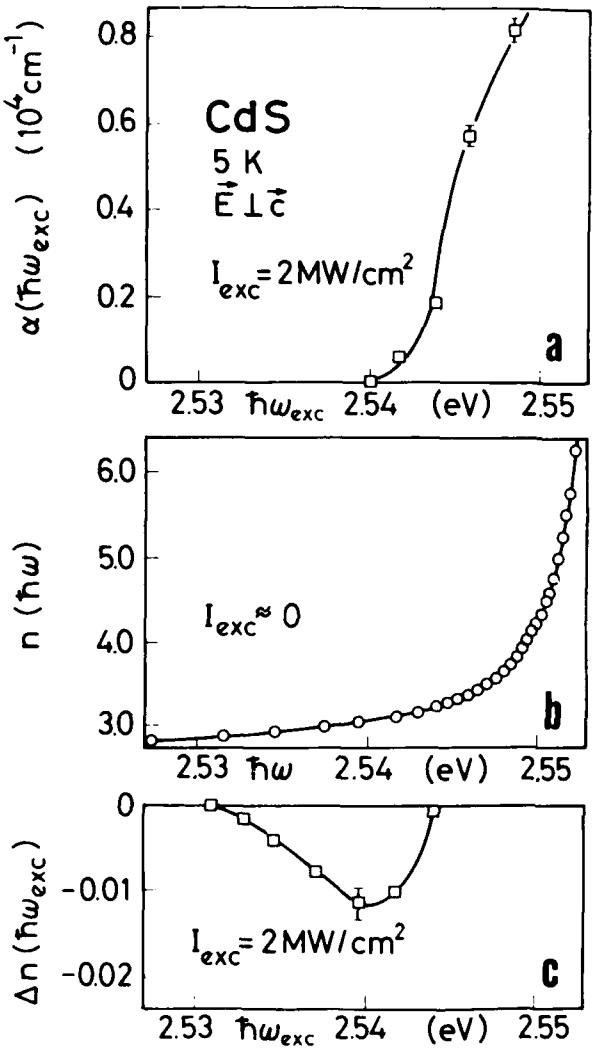


Fig. 7 : The absorption coefficient of a highly excited CdS sample at the photon energy $\hbar\omega_{exc}$ of the exciting laser (a), the spectrum of the real part of the refractive index of an unexcited sample and its variation under strong excitation (c).

schematically shows the transmission of a lossless FP as a function of the phase-shift δ of one round trip

$$\delta = 4\pi \, d \, n(\hbar\omega_{exc}, I_{exc}) \lambda_{vac}^{-1} \qquad (19)$$

For constructive interference of the partial waves, reflected from the back and front surfaces (i.e. $\delta = m \cdot 2\pi$; m = o,1,2,3..) a rather high amplitude builds up in the resonator, resulting in high transmission and low reflection. For $\delta = (2m + 1)\pi$; (m = o,1,2..) the interference in the FP is mainly destructive, the field amplitude in the resonator is small, yielding low transmission and high reflection. We assume now that a light beam of temporally increasing intensity is sent onto the sample, with a wavelength fulfilling the condition for a minimum in the transmission of the FP (point A in Fig. 6,9). With increasing incident intensity only a small fraction of h is transmitted (Fig. 9). At a certain value of $I_o(t)$, the intensity in the resonator is sufficiently high as to alter n. Consequently δ is changing (Eq. (19)), the interference in the FP is getting more constructive and one moves towards point B. Beyond B in the cavity a positive feedback sets in, between the light intensity in the FP and the change of n. This gives rise to a jump from point B to rather high transmission (point C in Fig. 8,9). A further increase of $I_o(t)$ shifts the "working point" in the FP towards the next minimum. If $I_o(t)$ is now increased, one moves back to point C. At further decrease of $I_o(t)$ the increasing constructive interference in the cavity maintains the high transmission level. One continues smoothly towards D and only from there the transition occurs to the low-transmission state again. As a consequence, one gets a hysteresis-loop, which is revolved counter clockwise (Fig. 9). For a certain range of I_o one hase two stable states, one of low and one of high transmission. The first occurs if one comes from low values of I_o, the latter if one approaches the bistable region from the high intensity side. Since we assumed a lossless FP, the intensity of the reflected beam exhibits just the opposite behaviour.
A similar hysteresis-loop as shown in Fig. 9 can also be obtained in transmission from a FP filled with an absorbing material ($\alpha^{-1} \ll d$) the absorption of which is bleached at high levels of the light intensities. In this case λ_{vac}(or $\hbar\omega_{exc}$) is selected in a way, that

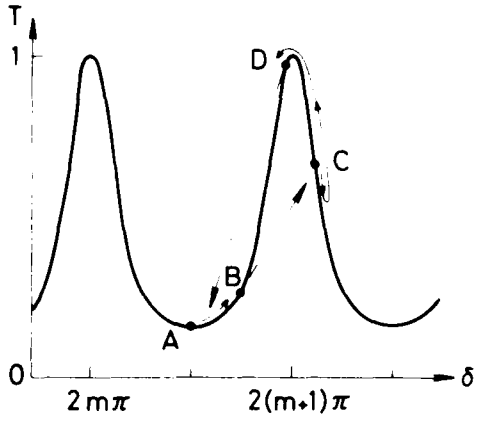


Fig. 8: Schematic representation of the transmission I through a Fabry Perot resonator as a function of the phaseshift δ of one round trip. The points A to D explain the appearance of dispersive optical bistability.

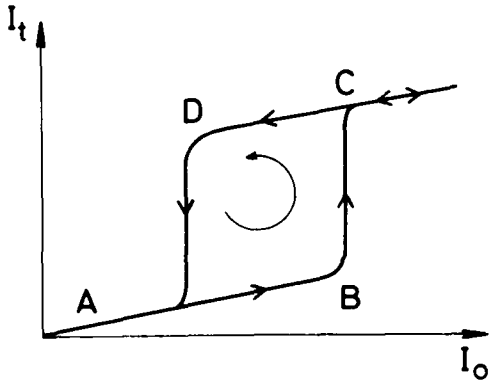


Fig. 9: The hysteresis loop connected with dispersive optical bistability.

constructive interference would occur, if the sample would be transparent. Because of the absorption no standing wave pattern is formed in the FP. With increasing I_o α decreases, constructive interference becomes possible, the light intensity in the FP increases even more resulting in further bleading of α. As a consequence of this positive feedback one gets again a step from low to high transmission. With decreasing I_o the constructive interference in the FP keeps α at low values down to a certain critical value of the incident intensity which is below the one necessary for the onset of bleading. Again a OB and a hysteresis like in Fig. 9 are observed in transmission. For more detailed theoretical investigation about these types of OB the reader is referred e.g. to the proceedings of three recent meetings on this subject [22-24]. Experimentally, dispersive OB partly connected with a decrease of α at $\hbar\omega_{exc}$ has been found e.g. in InSb, GaAs and InAs [8,22-27]. Dispersive OB connected with the transition to the biexciton has been predicted by [8,28] and observed in [29]. First indications of a dispersive bistability in CdS have been reported in [21,30] for photon-energies just below 2.54 eV (see Fig. 7c). Experiments with improved temporal resolution (see below) confirm this phenomenon. A typical hysteresis loop is shown in Fig. 10. Optical hysteresis in connection with nonlinear frequency mixing in CdS has been reported in [31].

In the following a type of OB is presented, which is connected with an increase of the absorption coefficient with increasing excitation. This idea has been presented together with experimental data to our knowledge for the first time by two of the authors (K.B. and C.K.) at the spring meeting of the German Physical Society in March 1983 and has subsequently been published in [30]. Later on, detailed theoretical investigation have been performed [32,33,34], the temporal resolution of the experimental setup has been improved [20,21] and it has been shown that this process is of rather general nature [35]. We shall give here a qualitative description, using the transition from a low density exciton gas to the EHP: The photon energy of the laser is chosen to be situated spectrally in a region, where the sample is rather transparent at low excitation and becomes absorbing at high excitation. According to Figs. 6 and 7a this is e.g. the region between 2.54 eV and the free exciton absorption around 2.552 eV. For weak incident intensities

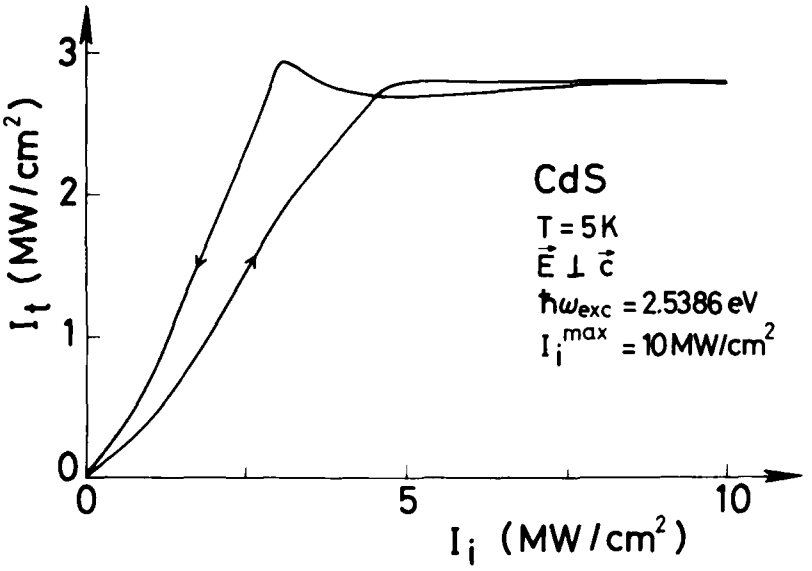


Fig. 10: Measured hysteresis loop of the dispersive optical bistability in CdS.

the sample is transparent. With increasing incident intensity electron-hole pairs are created either due to two-photon absorption or due to one photon absorption in the tail of the absorption edge [21]. If the density of carriers increases, the absorption edge is shifting to lower photon energies e.g. due to the renormalization of the gap connected with the formation of an EHP. Fig. 7 and the sample becomes opaque (Fig. 11). A further increase of I_{exc} keeps the sample in the state of low transmission, except that the density and thus the chemical potential of the plasma are raised to $\mu(n_p, T_p) = \hbar\omega_{exc}$. In this case the sample becomes transparent again (dashed line in Fig. 11) and phenomena like self-pulsation of the transmitted light may occur. If the sample is in the state of strong absorption, the excitation intensity may be lowered to values which are far below those necessary for the switch down and the sample still remains opaque, because the EHP is now pumped very efficient by strong one photon excitation. If the excitation level is reduced even so far that this process is no more sufficient to sustain the plasma, the

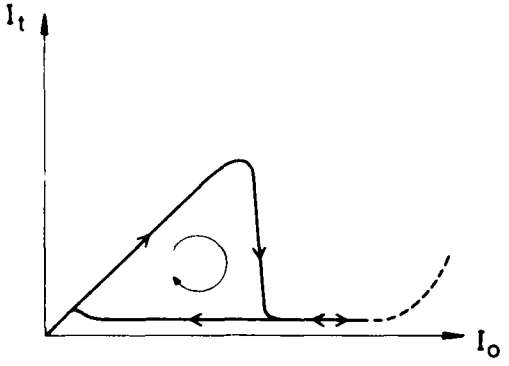


Fig. 11: Schematic drawing of the hysteresis-loop connected with an absorptive optical bistability.

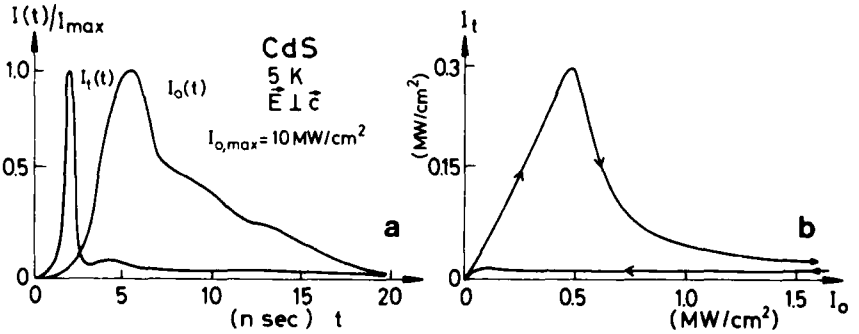


Fig. 12: Observed temporal evolution of the incident and the pulse transmitted through a thin CdS sample (a) and the resulting hysteresis loop (b).

carriers recombine, the absorption edge shifts again to higher photonenergies and the sample becomes transparent again. As a consequence one expects a hysteresis loop in the transmission, which is revolved clockwise in contrast to the dispersive one. Fig. 12 gives experimental results for CdS. The pulses of an excimer-laser pumped dye-laser are registered before and after the transmission through a CdS-sample with a streak-camera coupled to an optical multichannel analyzer. The incident laser passes over a grating in order to remove the super-radiant background emission from the dye-laser. The transmitted pulse passes also through a spectrometer to suppress contributions from the luminescence. The switching down is directly seen. The analysis of the switch-down time from single shot experiments gave values of about 65 psec. The recovery time it takes the sample to get transparent again seems to be considerably longer. In agreement with Fig. 7 this OB can be observed down to photon energies of about 2.54 eV [3o]. We call this OB an intrinsic one because it needs no external feedback from a FP. The feedback occurs in the electronic system itself. The increase of n_p causes an increase of the absorption via band renormalization effects. It should be noted, that this OB connected with the increase of absorption does not only work in connection with an EHP. An increase of the lattice temperature with excitation, which in turn causes a red-shift of the absorptionspectrum, works equally well. This has been demonstrated with multiple quantum well structures [35], and with a bound exciton absorption line in CdS [36]. The OB produced by an increase of the absorption is easier to handle than those involving a FP: The demands to the monochromacy and coherence length of the light and to the flatness and parallelism of the sample surfaces are much lower. CdS has so far a unique position as several different types of OB have been found in this material, namely the OB connected with degenerate wave mixing [31] the dispersive and absorptive OB connected with the formation of a EHP [3o], and the absorptive one connected with lattice heating and a shift of a bound exciton absorption line [36].

4. LOGIC ELEMENTS IN SEMICONDUCTORS REALIZED BY OPTICAL BISTABILITY

The investigation of OB is of considerable interest from various reasons. From the viewpoint of fundamental research OB allows e.g. to study nonlinear dynamics of the electronic system of semiconductors and phase transitions far away from equilibrium states. Concerning applied research, one has realized that OB-devices exhibit characteristica in the transmission and/or reflection of light beams, which are almost identical to those of the components of presently used computers with respect to the electric current. This will be shown in the following (see also [37]).

Every optically bistable device may be considered as an optical memory. The states of low and high transmittivity (or reflectivity) can be identified with the logical zero and one. If the device is in one stable state, it may be brought to the other by temporarily increasing or decreasing the incident intensity. (Fig. 13a, b).

If the bistable device is constructed in a way that the width of the hysteresis-loop is made small, it can be used as logic gate. A curve like the one shown in Fig. 13c, which is characteristic for dispersive OB or bleaching of absorption in a FP (see above), acts as a AND-gate if the switching point is situated between once and twice the intensity defining the logic one (Fig. 13d). If the switching occurs between intensities representing the upper limit for zero and the lower limit for one one has a OR-gate. (Fig. 13c). Transmission characteristics which are due to a dispersive OB in reflection or the absorptive bistability discussed above can be regarded as an inverter or NOT-gate, because they have high transmittivity at low incident intensities and low transmittivity at high intensities. Fig. 13 e+f. The diffracted orders ± 1, ± 2 ... from a LIG also are forming an AND-gate. They occur only if the beams "a" and "b" are falling simultaneously onto the sample. Some of the diffracted beams of a LIG are phase-conjugate with respect to one of the incident beams. This phenomenon gives very powerful possibilities of wavefront reconstruction. This aspect is however beyond the scope of this article. The reader is referred e.g. to [1o] and the references therein.

Memories, and the three different gates mentioned above are the basic constituents necessary to construct a computer. Since all of these functions can be realized optically there is a lot of speculation and enthusiasm about all optical computers, handling data as light pulses and not as electric current or voltage signals. Since some of the nonlinear phenomena presented in chapter two may have time constants of the order of $1o^{-12}$sec (e.g. a virtual creation of biexcitons [28] such optical computers would be much faster than even the best presently available electronic ones, with time-constants around $1o^{-9}$sec. Parallel processing would be another advantage of such optical systems. Since light beams can principally be focussed down to a spot of a diameter roughly equal to the wavelength, i.e. to about $1\mu m^2$ a chip of $1\,cm^2$ allows to handle in parallel $1o^8$ beams. With a processing time constant of $1o^{-11}$sec this would result in $1o^{19}$ operations per cm^2 and sec. Low switching energies down to $1o^{-13}$ Ws/μm^2 have been reported [35]. Finally it should be mentioned that one sample may simultaneously used for various logic connections.

Though these properties are fascinating and promissing for future technical applications there are still many difficulties to overcome before the first all optical computer works. Some of them are the following: The OB devices with low switching energies which easily allow CW operation are generally rather slow with switching times in the range from tens of nsec to msec [35,36]. This is trivial for all mechanisms which depend on thermal effects. Electronic excitations in narrow gap semiconductors like InSb, InAs, GaAs or CdHgTe have long lifetimes due the decrease of the recombination probability with transition energy. There are hopes to reduce the lifetime by heavy doping, but it is not yet clear, how this will influence the nonlinear properties.

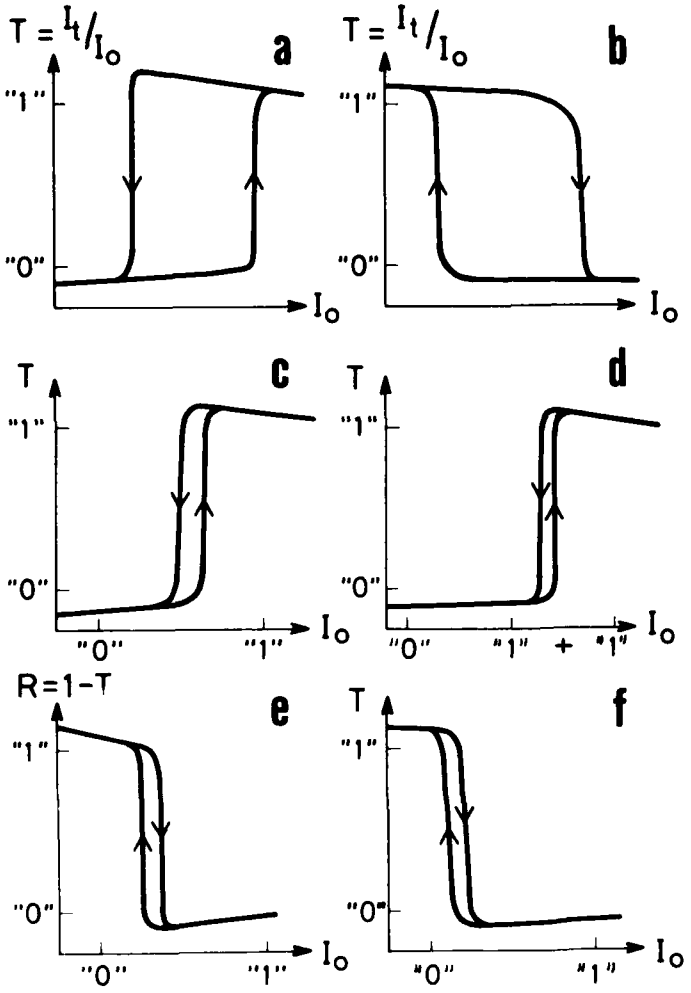


Fig. 13: Schematic representations of logic connections that may be realized by nonlinear optical devices. Memories (a) and (b); "OR" (c), "AND"(d), "NOT" or inverter (e) and (f)

Processes which are fast e.g. the virtual excitation of biexcitons in CuCl or the creation of a plasma in CdS lead to OB only under excitation intensities in the MW/cm^2 range which can be applied to the samples only for tens of nsec. Furthermore several of the OB devices work at liquid He or N_2 temperatures only.
Another problem is, that the output of an optical gate is generally considerably lower than the input intensity. It is therefore difficult, to drive with the output several other gates and consecutive operation of several logic gates is presently almost impossible. There are semiconductors, which are able to amplify light under suitable excitation (see e.g. Fig. 6 [3,37]) but they again need usually cooling and pump intensities of tens of kW/cm^2 up to the MW/cm^2 range.

In spite of the above mentioned still unsolved problems, there is in our opinion a good chance to introduce digital optical data handling in connection with data transfer through fiber-optics. If data are transferred through fibers, necessarily electronic pulses have to be transformed into optic ones and vice-versa. The first step of optical data processing could consist in only a few logic connections at the input and output of the fibers. If they work satisfactorily, the interface between electronic and optic signals may be pushed further and further away from the entrance or exit of the fiber into the data-handling system itself. If the pure electronic part converges to zero, resulting in a real all optical computer, will be seen in the next decades.

Acknowledgements: The results shown here are part of a project of the Sonderforschungsbereich Festkörperspektroskopie, financed by the Deutsche Forschungsgemeinschaft. The high quality CdS samples have been grown in the Kristall- und Materiallabor der Universität Karlsruhe.

LITERATURE

[1] R.S. Knox, Theory of Excitons, Solid State Physics, supplement 5, F. Seitz and D. Turnbull, Academic Press (1963)

[2] Topics in Current Physics, Vol. 14, Excitons, editor: K. Cho, Springer (1979)

[3] C. Klingshirn and H. Haug, Phys. Reports 7o (1981) 315

[4] Modern Problems in Condensed Matter Sciences, Vol. 2 Excitons, eds.: E.I. Rashba and M.D. Sturge, North Holland (1982)

[5] Collective Excitations in Solids, editor: B. Di Bartolo, NATO ASI Series B Vol. 88, Plenum Press (1983)

[6] C. Klingshirn in Proceedings of the NATO ASI on Energy Transfer in Condensed Matter, Erice (1983) to be published, editor B. Di Bartolo

[7] H.J. Eichler, Advances in Solid State Physics, Vol. XVIII (1978) p 241

[8] A. Müller, D.A.B. Miller and S.D. Smith, Advances in Physics, 3o (1981) 697

[9] H. Haug, Advances in Solid State Physics XXII (1982) p 149

[1o] M. Ducloy, ibid. p 35

[11] H. Kalt, V.G. Lyssenko, R. Renner and C. Klingshirn, to be published in Sol. State Commun.

[12] V.G. Lyssenko, K. Kempf, K. Bohnert, G. Schmieder, C. Klingshirn and S. Schmitt-Rink, Sol. State Commun. 42 (1982) 4o1

[13] H. Haug and S. Schmitt-Rink, to be published in Progress in Quantum Electronics, J.H. Sanders and S. Stenholm eds.

[14] M. Rösler and R. Zimmermann, phys. stat. sol. b83 (1977) 85

[15] G. Beni and T.M. Rice, Phys. Rev. B18 (1978) 768

[16] K. Bohnert, M. Anselment, G. Kobbe, C. Klingshirn, H. Haug, S.W. Koch, S. Schmitt-Rink and F.F. Abraham, Z. Physik B42 (1981) 1

[17] T.M. Rice, Solid State Physics, Vol. 32, Academic Press (1977) p 1, J.C. Hensel, T.G. Philips and G.A. Thomas ibid. p 88

[18] K. Kempf, Ph.D.Thesis, Frankfurt (1984) partly published in K. Kempf and C. Klingshirn, Sol. State Commun. 49 (1984) 23

[19] A. Kreissl, K. Bohnert, V.G. Lyssenko and C. Klingshirn, phys. stat. sol.b 114 (1982) 537

[2o] C. Klingshirn, K. Bohnert, K. Kempf and H. Kalt in Ref [24]

[21] K. Klingshirn, K. Bohnert, H. Kalt, V.G. Lyssenko and K. Kempf, to be published in Proc. of the 3rd IUPAP Semiconductor Symposium on "High Excitation and Short Pulse Phenomena", Trieste (1984)

[22] Proc. of the Intern. Conf. on Optical Bistability, held in Asheville, N.C. (198o) Ch.M. Bowden, M. Cifton and H.R. Robl, eds.: Plenum Press (1981)

[23] Proc. of the Topical Meeting on Optical Bistability held in Rochester, N.Y. (1983) to be published

[24] Proc. of a Meeting for Discussion on "Optical Bistability, Dynamical Nonlinearity and Photonic Logic", to be published in Transactions of the Royal Society (1984)

[25] H.M. Gibbs, S.L. McCall, T.N.C. Venkatesan, A.C. Gossard, A. Passner and W. Wiegmann, Appl. Phys. Lett. 35 (1979) 451

[26] D.A.B. Miller, S.D. Smith and A.M. Johnston, Appl. Phys. Lett. 35 (1979) 658

[27] C.D. Poole and E. Garmire, Appl. Phys. Lett. 44 (1984) 363

[28] S.W. Koch and H. Haug, Phys. Rev. Lett. 46 (1981) 45o; E. Hanamura, Sol. State Commun. 38 (1981) 939

[29] B. Hönerlage, J.Y. Bigot and R. Levy in Ref [23], N. Peyghambarian, D. Sarid and H.M. Gibbs in Ref. [23], R. Levy, J.Y. Bigot, B. Hönerlage and J.B. Grun, Sol. State Commun. 48 (1983) 7o5, N. Peyghambarian, H.M. Gibbs, M.C. Rushford and. D.A. Weinberger, Phys. Rev. Lett. 51 (1983) 1692

[30] K. Bohnert, H. Kalt and C. Klingshirn, Appl. Phys. Lett. 43 (1983) 1088,
K. Bohnert in Ref. [6]

[31] A. Borshch, M. Brodin, V. Volkov and N. Kukhtarev, Optics Commun. 41 (1982) 213

[32] H.E. Schmidt, H. Haug and S.W. Koch, Appl. Phys. Lett. 44 (1984) 787

[33] F. Henneberger and H. Roßmann, phys. stat. sol. b121 (1984)

[34] See the contributions of H. Haug, of S.W. Koch and of F. Henneberger to the 3rd IUPAP Semiconductor Symposium on "High Excitation and Short Pulse Phenomena", Trieste (1984)

[35] D.A.B. Miller, Ref. [24] and
D.A.B. Miller, A.C. Gossard and W. Wiegmann to be published in Optics Letters (1984)

[36] M. Dagenais in Ref. [24] and M. Dagenais and W.F. Shorfin, to be published in Appl. Phys. Lett.

[37] E. Abraham, C.T. Seaton and S.D. Smith, Scientific American 248 (1983) Nr. 2 p 63

DISCUSSION

S.D.Smith, UK

Were your measurements made with pulses?

Author's Reply

This is one of the problems which I addressed. The devices which work continuous with lasers and which have very low switching powers in the milliwatt area until now at least tend to be rather slow with switching times in the 10 to 100 nanosecond region or even longer. The physical mechanisms which promise switchings with picoseconds generally need rather high intensities. For example, if you go to, say, biexcitons in copper chloride, Hanamura has shown that you can, from a theoretical point of view, expect picosecond switching times, but you need then megawatts per square centimetres and with this you can not apply a continuously working beam on the sample — you can apply it only in short pulses.

H.Gibbs, US

To see increasing absorption bistability you used band-gap renormalization. In many systems the band fills up as fast as the edge shifts down, so one sees no shift of the actual absorption edge to lower energies. Why is it that in cadmium sulphide this red shift is so pronounced?

Author's Reply

There is a competition between the decrease of the gap and the filling of the states with increasing electron-hole pair density. It depends on the material parameters, which term is dominating. In CdS and many other semiconductors there is a certain range of densities, where the electron-hole plasma is bound with respect to the free exciton, i.e. there is a red shift of the absorption-edge. In any case, both contributions have to be considered in a quantitative description of the variations of the optical properties in the transition from low densities to the plasma.

J.W.Haus, US

Has optical bistability been observed with the exciton-biexciton mechanism in cadmium sulphide and if not, why not?

Author's Reply

The point is the following: In CdS, if you want to produce so many excitons that you get a measurable variation of the refractive index you are already at the point where you produce an electron-hole plasma. The situation in copper chloride is much more favourable. In copper chloride it is more favourable to use nonlinearities connected with the biexciton to get optical bistability, in cadmium sulphide the nonlinearity is dominated by the electron-hole plasma.

J.W.Haus, US

So it's a matter of the binding energies?

Author's Reply

It's a matter of the binding energies of the size of the exciton, the screening length, and all these things, yes.

Unidentified Speaker

You mentioned a lot of materials and you mentioned the bleaching of absorption mechanism. The biological eye operates on that mechanism also, by way of something that's called rhodopsin and which bleaches one molecule for every two quanta of light. It may be a lot slower, but it's very sensitive. Have you considered this?

Author's Reply

No. I was only discussing semiconductor materials, not biological systems. But it is interesting to look at these, especially if they have such high sensitivity.

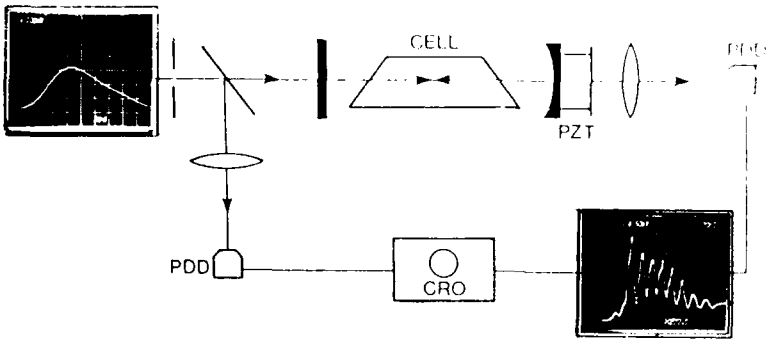


Fig. 2. Experimental arrangement of Fabry-Perot system and signal detection.

total resolution $\leq$ 1 ns. Intra-cavity gas cells were normally terminated with Brewster angled windows to minimise losses although compact systems sealed by cavity optics are envisaged for device applications. The cavity was pumped by a transversely excited atmospheric (TEA) CO_2 laser operating on a single mode giving temporally smooth pulses (duration $\sim$ 100 ns) with peak powers $\sim$ 1 MW. The transverse intensity profile of the input signal in Gaussian with 1/e spot diameter ca. 3 mm. An optical delay line prevented feedback of the laser pulses from the nonlinear resonator to the laser.

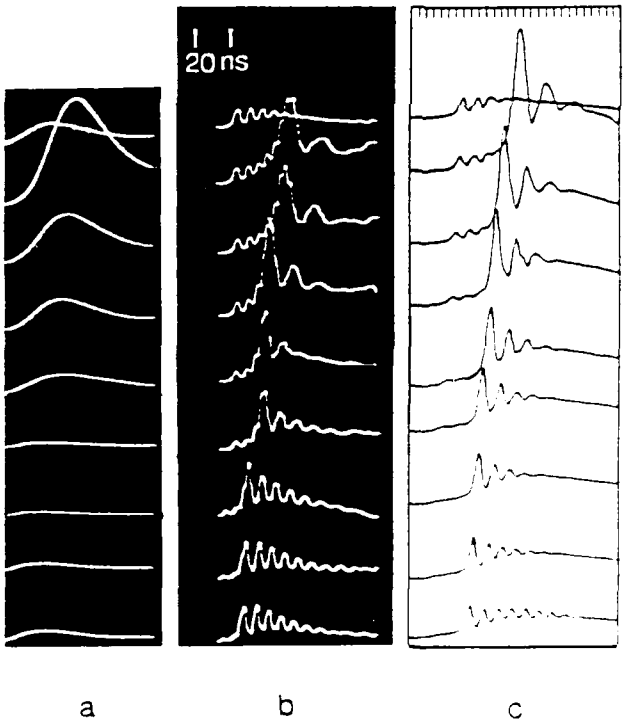


Fig. 3. PZT scan of a) empty cavity, b) cavity with 1 Torr SF_6 in 91 cm cell; input and output reflectivities 67% and 36% respectively, c) computer traces for parameters corresponding to b); $\alpha L = 1.5$; $\Delta = 1$; tick spacing t_R.

Theoretical Considerations

Routes to chaos in nonlinear systems have attracted a great deal of beautiful, but rather abstract, theory. However the fundamental "two-round trip time" modulation, which is of primary interest here, can be given a neat physical interpretation (9). The $\chi^{(3)}$ responsible for bistability also induces four-wave mixing, so that a "signal" field Stokes-shifted from a strong pump field will experience parametric gain (in conjunction with its anti-Stokes field) at the expense of the pump field. If we place the system in a resonator then the feedback allows spontaneous oscillation if the gain is large enough. The most favorable case has both Stokes and anti-Stokes fields cavity resonant; the nice thing about the present type of nonlinear-mixing is that such a double resonance can be guaranteed, due to the nonlinear index shift: as the pump intensity is increased, the "dressed" cavity modes "transphase" in frequency until they straddle the pump frequency (Fig. 1). This is the double resonance position and if the gain exceeds the cavity losses, the system bursts into oscillation: the beat frequency between the sidebands and the pump is then just $(2t_R)^{-1}$. It can be shown that the above argument is quantitatively identical to more formal treatments of the Ikeda instability (2,10).

This discussion has ignored the necessarily finite bandwidth of the nonlinear response: clearly if the bandwidth is much smaller than t_R^{-1} there will not be enough gain in the doubly-resonant configuration: conversely if the bandwidth is much greater than t_R^{-1}, further pairs of sidebands can break into oscillation, leading to complex waveforms. This demonstrates the desirability of being able to tailor the bandwidth to be of order t_R^{-1}, as is possible in gases.

The chaotic regime occurs when the parametric four-wave gain gets too far above threshold: this can be controlled by appropriate choice of cavity parameters.

As a final remark on this model, note that it is couched in such general terms that one can expect phenomena of this type in a wide variety of systems: at least four have been demonstrated to date (1,4,5,7).

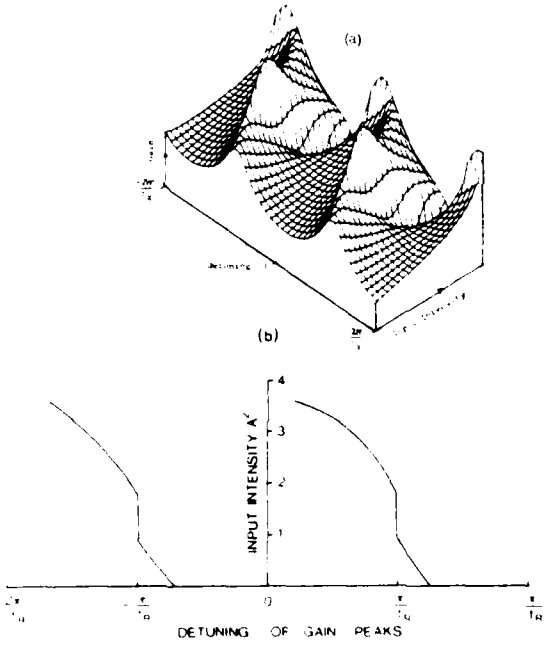


Fig. 1. Gain spectrum for a ring-cavity with $R = 0.58$, $\phi_0 = 2.3$. $|E_t|^2$ ranges from zero to just below the instability threshold. R is the amplitude feedback factor of the cavity and ϕ_0 the mis-tuning.

Experimental Considerations

While ring cavity systems may be easily analysed, involving only unidirectional propagation, they do not readily fulfil practical requirements of compact design and ease of operation necessary for device applications. Our recent experimental investigations have therefore been directed to Fabry-Perot systems in which, as we have previously predicted (9) period-doubling cascades to chaos may be generated. A typical laboratory arrangement is illustrated in Fig. 2. The cavity comprises single surface partially transmitting germanium optics suitably coated for 10 μm CO_2 laser radiation, one provided with piezo-electric transducer (PZT) for fine cavity length control. Resonator lengths were varied from 2 - 15 cm. The input and output signals were monitored by photon drag detectors and a Tektronix 7104 oscilloscope,

OPTICAL MODULATORS AND BISTABLE DEVICES USING MOLECULAR GASES

R.G. Harrison, W.J. Firth, I.A. Al-Saidi and E. Cummins

Department of Physics, Heriot-Watt University, Riccarton, Edinburgh EH14 4AS, U.K.

Summary

We have predicted that so-called 'bistable' nonlinear Fabry-Perot resonators can actually exhibit a rich variety of instabilities: in particular the resonator can convert a steady input field into an oscillatory or even chaotic output field. Oscillation itself leads to a passive all optical laser modulator device tunable in frequency in the range ~ 1 GHz with anticipated applications in optoelectronics and telecommunications. The unique advantages of molecular gases as nonlinear media in these systems has been utilised in successfully demonstrating these effects using pulsed CO_2 laser radiation, the results of which are in excellent accord with theory, analysis also showing that prospects for cw operation are promising.

Introduction

In many optical processing systems there is a need for amplitude- or frequency-modulation of optical beams. Presently-available schemes may be active, directly modulating the laser output, or passive, taking a cw laser beam and applying, e.g., electro-optic or acousto-optic modulation to it. In this paper we discuss a new passive modulation technique based on nonlinear refractive index effects, which has potentially significant advantages over other techniques, and we present some experimental demonstrations of the principle. The proposed device could amplitude-modulate smooth pulses at tunable frequencies in the range ~ 1 GHz. It would be passive; it could modulate any laser for which a suitable medium could be found; and it could transfer modulation from one laser to another as a dynamic "transphasor" (1). Applications already exist for such a device in "sideband spectroscopy" and we anticipate applications in optoelectronics and telecommunications.

This possibility arises as a by-product of substantial effort over the last few years on period-doubling cascades to chaotic behaviour (1,2,3,4) in optically bistable systems. Ikeda (2) showed in 1979 that an optically bistable ring resonator containing a two-level system can show a period-doubling cascade, a sufficiently strong cw input beam yielding an output oscillating at twice the resonator round trip time t_R. On further increasing the input field the output period doubles to chaos.

While the chaos itself has few obvious applications, it is important both in confirming the physical models and as a regime of operation to be avoided: in fact, as will emerge below, it is possible, by appropriate choice of parameters, to favour oscillation over chaos.

Nonlinear intensity dependent refraction, which lies behind the predicted phenomena, first appears to order $\chi^{(3)}$. The nonlinearity can be non-resonant, in which case it will be fast ($\tau < 1$ ps) but weak; a liquid or solid medium and high-power laser is then indicated for observation of dynamic instabilities. More likely, however, to discrimination against unwanted competing processes such as Raman or Brillouin scattering and/or optical damage. A resonant nonlinearity, on the other hand, can give bistability at very low powers (mW), but at the expense of response times (τ) usually too long for useful modulation. Liquid crystals and amorphous semiconductor films also show large $\chi^{(3)}$, associated with structural changes: again the response time is too long for our purposes.

Ideally, then, we would like a $\chi^{(3)}$, enhanced sufficiently to give low thresholds and outweigh competing processes, but with a response time of a few nanoseconds to allow a reasonably compact resonator design and direct observation of oscillatory transmission modulation. Gases are the only immediately available media with these properties: we are further led to look at infrared transitions, because then pressure variation allows us to vary τ through the target range 1 - 1 ns (in the visible, Doppler-broadening would be an undesirable complication). Such considerations enabled us to achieve the first observation of such effects in passive, all optical systems containing 2-level media; NH_3 gas contained in a ring cavity and pumped by the smooth 100 ns pulses from a TEA CO_2 laser (5,6). We identified and demonstrated the unique advantages of molecular gases in this field, providing in particular a vast range of laser-molecular coincidences together with control of response time, and therefore switching time and modulation frequency, by varying gas pressure.

Here we describe extension of this work to all optical Fabry-Perot resonators containing NH_3 and other gases providing compact systems which are shown to exhibit huge effects (7). Deep and sustained oscillation at twice the round trip time of the cavity ($2t_R$ at present ~ 150 MHz) along with t_R and t_R modulation and generation of chaos in the transmitted signal have all been obtained in excellent agreement with theoretical predictions. Furthermore, under other parameter conditions a rich variety of "optically bistable" effects are observed including switching, overshoot and power limiting; obtained in systems as small as 1 mm (8).

Already of value as a compact and simple laboratory device for temporal control and modulation of high power pulsed infrared radiation our analysis show the prospects for cw operation are very promising.

27. J.A. Mroczkowski, D.A. Nelson, "Optical absorption below the absorption edge in $Hg_{1-x}Cd_xTe$", J. Appl. Phys., Vol. 54, 1983. pp 2041-2051.

ACKNOWLEDGEMENT

We are grateful to Mullard Ltd for providing the samples and A.B. Dean for sample preparation.

DISCUSSION

J.W.Haus, US

I'd like to ask about the theory of the effects of the overshoot as a function of time. That could be due to some three dimensional effect. Does your theory have plane-wave propagation?

Author's Reply

We only have the carrier equations for the material itself in there, coupled with the Fabry-Perot equation. It's quite a simple model. We're using plane waves because we've made some effort to eliminate effects due to transverse structure of the beam in the experiments.

H.Gibbs, US

Would you elaborate on the inhomogenities and the evidence for them?

Author's Reply

The band tail absorption in CdHgTe extends a long way from the band gap. This is probably due to local fluctuations inthe band gap because of the random distribution of the alloy atoms. Bulk samples also show gradients in alloy composition both along the boule and radially. Within our typical spot diameter of about 100 microns we would expect significant refractive index variations. Scanning a spot across the sample usually gives a variation of at least one Fabry-Perot cavity order over a few mm but not as a smooth variation.

S.D.Smith, UK

What are the prospects for using CdHgTe at much larger band gaps? There is, of course, not only interest at 10 microns but at 5, maybe 3, maybe 1.5. I'd like to know how the material changes as you go to these conditions. One comment: The difference between InSb and CdHgTe is in the relative signs of thermal effects versus electronic effects. We have some new ideas to be talked about later in which we found it an interesting idea to have them (the thermal effect and the electronic effect) in the same direction. And so my first question, perhaps, has more point in this context.

Author's Reply

Everything we've done seems to agree very well with the analysis using the band filling nonlinearity plus free carrier. The band gap dependencies do agree with theory. So we could confidently do the scaling, I think, to smaller band gaps. The material itself is quite amenable to being grown with larger band gaps, as long as one can convince somebody to grow it. I mean the main advantage of using this material is, of course, that it's an important infrared detector material; that's why our material is grown with these small band gaps. But I think it would be interesting to look at larger band gaps. The main problem is material quality — getting it homogeneous enough.

REFERENCES

1. A. Miller, D.A.B. Miller, S.D. Smith, "Dynamic non-linear optical processes in semiconductors", Adv. Phys., Vol. 30, 1981, pp 697-800.

2. C.M. Bowden, H.M. Gibbs, S.L. McCall, Eds., "Optical Bistability, 2", New York, Plenum, 1984.

3. "Optical bistability, dynamic nonlinearity and photonic logic", Phil. Trans. Roy. Soc. London - Proc. Discussion Meeting, London, 1984, in press.

4. R. Dornhaus, G. Nimtz, "The properties and applications of the $Hg_{1-x}Cd_xTe$ alloy system", Springer Tracts in Modern Physics, Vol. 78, G. Hohler, Ed., Berlin, Springer-Verlag, 1976.

5. J.R. Hill, G. Parry, A. Miller, "Nonlinear refractive index changes in CdHgTe at 175K with 10.6 μm radiation", Opt. Commun., Vol. 43, 1982, pp 151-156.

6. G. Parry, A. Miller, R. Daley, "Nonlinear Fabry-Perot transmission in a CdHgTe etalon" in "Optical Bistability, 2", C.M. Bowden, H.M. Gibbs, S.L. McCall, Eds., New York, Plenum, 1984, pp 289-296.

7. A. Miller, G. Parry, R. Daley, "Low power nonlinear Fabry-Perot reflection in CdHgTe at 10 μm", IEEE J. Quantum Electron. Vol. Qe-20, 1984, pp 710-715.

8. J.G.H. Mathew, D. Craig, A. Miller, "Optical switching in a CdHgTe etalon at room temperature", Appl. Phys. Lett., 1984, in press.

9. D. Craig, A. Miller, J.G.H. Mathew, A.K. Kar, "Fast optical switching and bistability in room temperature CdHgTe at 10.6 μm", Infrared Phys., 1984, in press.

10. G.L. Hansen, J.L. Schmit, T.N. Casselman, "Energy gap versus alloy composition and temperature in $Hg_{1-x}Cd_xTe$", J. Appl. Phys., Vol. 53, 1982, pp 7099-7101.

11. B.S. Wherrett, N.A. Higgins, "Theory of nonlinear refraction near the band edge of a semiconductor", Proc. Roy. Soc. London, Vol. A379, 1982, pp 67-90.

12. B.S. Wherrett, "A comparison of theories of resonant nonlinear refraction in semiconductors", Proc. Roy. Soc. London, Vol. A390, 1983, pp 373-396.

13. R. Jain, M.B. Klein, "Degenerate four-wave mixing near the band gap of semiconductors", Appl. Phys. Lett. Vol. 35, 1979, pp 454-456.

14. D.A.B. Miller, C.T. Seaton, M.E. Prise, S.D. Smith, "Band gap resonant nonlinear refraction in III-V semiconductors", Phys. Rev. Lett., Vol 47, 1981, pp 197-200.

15. A. Miller, G. Parry, "Optical bistability in semiconductors with density dependent carrier lifetimes", Opt. Quantum Electron., Vol. 16, 1984, pp 339-348.

16. J. Blakemore, "Semiconductor Statistics", Oxford, Pergamon, 1962, Ch.6.

17. P.E. Peterson, "Auger recombination in mercury cadmium telluride", Semiconductors and Semimetals, Vol. 18, New York, Academic Press, 1981, Ch.4.

18. D.A.B. Miller, M.H. Mozolowski, A. Miller, S.D. Smith, "Nonlinear optical effects in InSb with a cw CO_2 laser", Opt. Commun., Vol. 27, 1978, pp 133-136.

19. D. Weaire, B.S. Wherrett, D.A.B. Miller, S.D. Smith, "Effect of low-power nonlinear refraction on laser-beam propagation in InSb", Opt. Lett., Vol. 4, 1979, pp 331-333.

20. D.A.B. Miller, "Refractive Fabry-Perot bistability with linear absorption: Theory of operation and cavity optimisation", IEEE J. Quantum Electron., Vol. QE-17, 1981, pp 306-311.

21. C.R. Pidgeon, B.S. Wherrett, A.M. Johnston, J. Dempsey, A. Miller, "Two-photon absorption in zinc-blende semiconductors", Phys. Rev. Lett., Vol. 42, 1979, pp 1785-1788.

22. M.H. Weiler, "Nonparabolicity and exciton effects in two-photon absorption in zinc-blende semiconductors", Solid State Commun., Vol. 39, 1981, pp 937-940.

23. A. Miller, A. Johnston, J. Dempsey, J. Smith, C.R. Pidgeon, G.D. Holah, "Two-photon absorption in InSb and $Hg_{1-x}Cd_xTe$", J. Phys. C., Vol. 12, 1979, pp 4839-4849.

24. A. Johnston, C.R. Pidgeon, J. Dempsey, "Frequency dependence of two-photon absorption in InSb and $Hg_{1-x}Cd_xTe$", Phys. Rev. B, Vol. 22, 1980, pp 825-831.

25. T. Bischofberger, Y.R. Shen, "Theoretical and experimental study of the dynamic behaviour of a nonlinear Fabry-Perot interferrometer", Phys. Rev. A., Vol. 19, 1979, pp 1169-1176.

26. E. Finkman, Y. Nemirovsky, "Infrared optical absorption of $Hg_{1-x}Cd_xTe$", J. Appl. Phys., Vol. 50, 1979, pp 4356-4361.

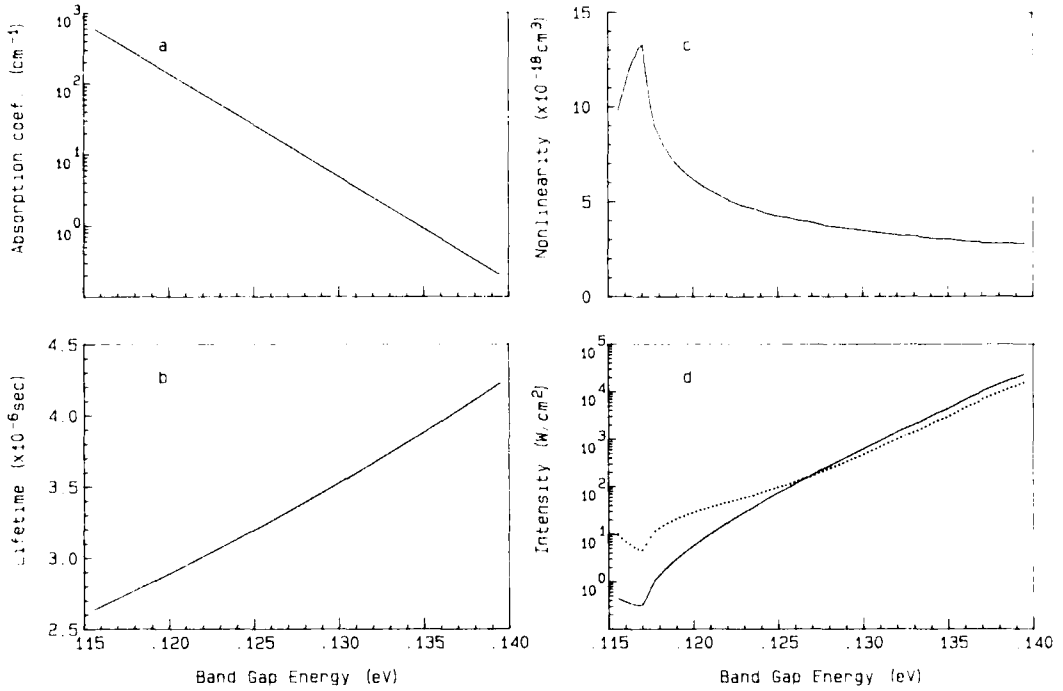


FIGURE 8. Band gap energy dependences of the properties of a CdHgTe etalon at 77K as discussed in the text; (a) absorption coefficient, (b) Auger recombination lifetime, (c) optical nonlinearity, σ, and (d) solid line: mean cavity intensity, dotted line: incident intensity.

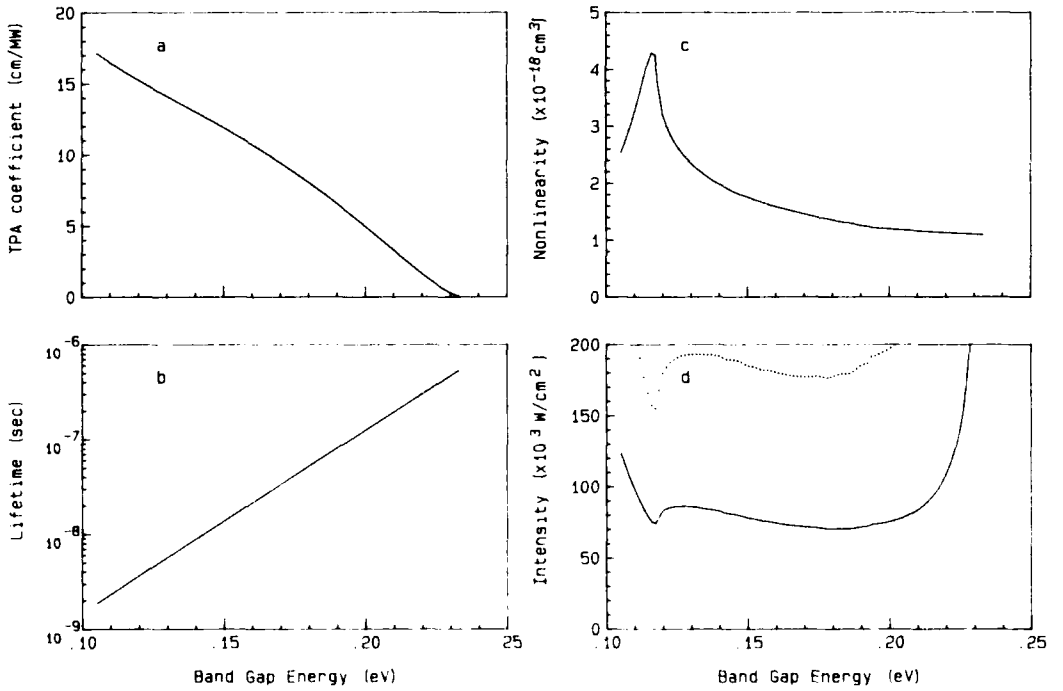


FIGURE 9. Band gap dependences of the properties of a CdHgTe at 295K as discussed in the text; (a) two-photon absorption coefficient, (b) Auger recombination lifetime, (c) optical nonlinearity, , and (d) solid line: mean cavity intensity, dotted line: incident intensity.

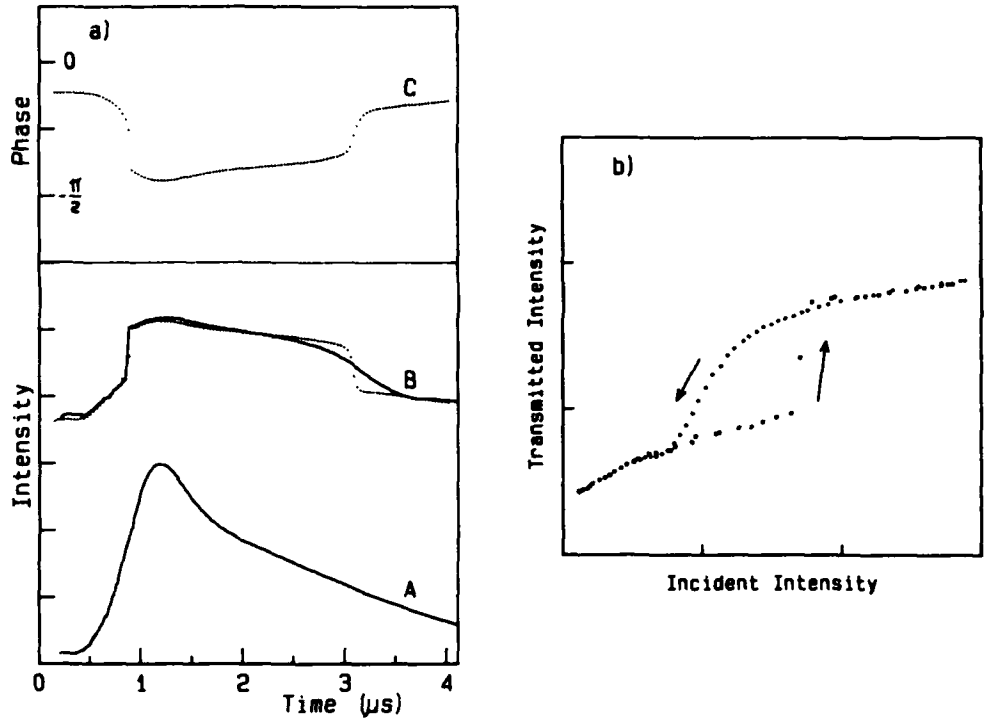


FIGURE 7(a). Pulse shapes, A: incident 1.75 μs (fwhm), B: transmitted, for a 363 μm thick $Cd_{0.23}Hg_{0.77}Te$ etalon at 295K. Solid lines are experimental; dotted lines are theory. Peak incident intensity 500 kW/cm^2. C: calculated etalon phase.
(b) Incident versus transmitted intensity for the experimental results.

$$\alpha = \alpha_o \exp\left[\frac{\sigma(E-E_o)}{T+T_o}\right]$$

for $20 \leq \alpha\ (cm^{-1}) \leq 1000$ where in $\alpha_o = -20.44 + 51.70x$, $\sigma = 5.65K/cm^{-1}$, $E_o = -3109 + 16450x$ (cm^{-1}) and $T_o = 80.5K$. It is convenient to make use of this for the present analysis and the resulting absorption coefficient at 10.6 μm (0.117 eV) as a function of band gap energy according to this expression and Eq.(1) is shown in Figure 8(a). (Note that free carrier absorption has been neglected here but will be a significant background contribution for larger band gaps.) Auger recombination at small excitations has been studied by several workers[17] and the expected lifetime dependence on band gap in the same range is shown in Figure 8(b). The nonlinearity in Figure 8(c) is calculated according to Eq.(2). In Figure 8(d), we show the mean cavity intensity at 10.6 μm required to cause a half-wave phase shift in a single pass of a 200 μm thick etalon from Eqs.(3),(4),(6) and (12). We also show the calculated incident intensity required to take an etalon with reflectivities of 32% from one transmission maximum to the next, Eq.(13). These results show the strong band gap resonance expected for single photon nonlinear Fabry-Perot effects of CdHgTe at 10.6 μm.

A similar analysis for room temperature two-photon excitation is shown in Figure 9. The theoretical two-photon absorption dependence has been used. The background absorption from free carrier absorption has been included using an absorption cross section[27] of 4.6×10^{-16} cm^2. We note that the two photon excitation process does not show a strong band gap dependence however single photon effects would have to be considered when the photon energy approaches the band gap.

6. CONCLUSIONS

CdHgTe exhibits one of the largest band gap resonant optical nonlinearities discovered to date. At low temperature, nonlinear Fabry-Perot effects can be observed at less than 1 mW incident power. This is significant for all-optical applications in the 10 μm region because of the importance and widespread use of CO_2 lasers. Inherent thermal problems in CdHgTe necessitates the use of pulsed radiation. Optical bistability can be observed in CdHgTe at room temperature making use of pulsed, two-photon excitation. The nonlinearities deduced from these experiments are in good agreement with the theory of band gap resonant effects.

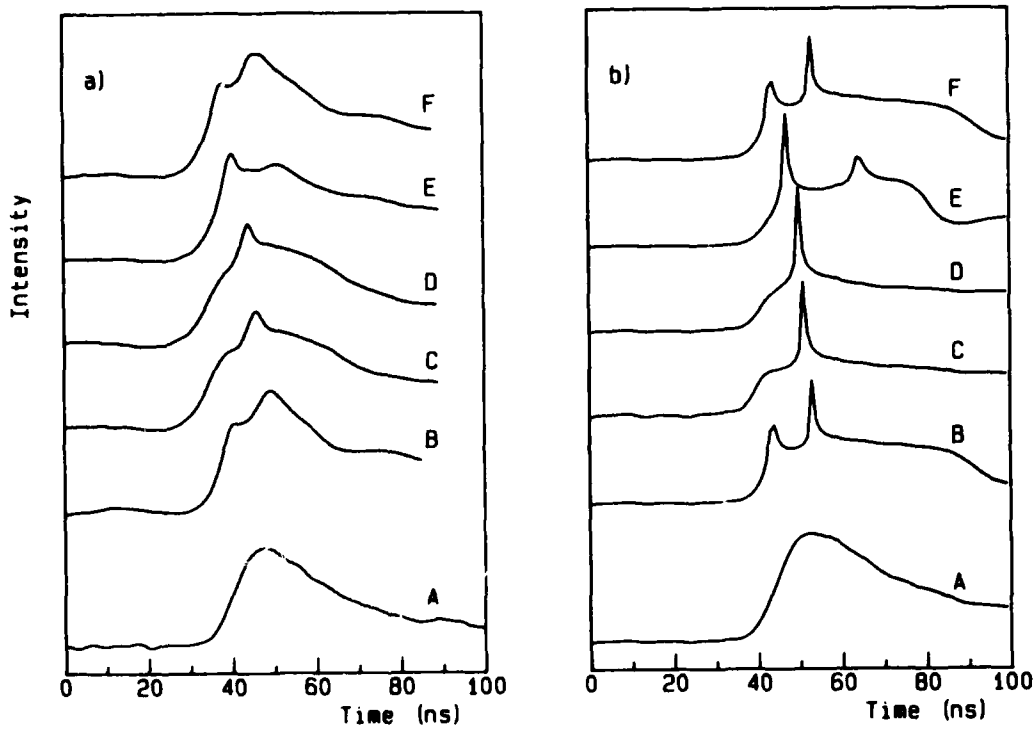


Figure 6 (a). Measured transmitted 30 ns pulse shapes for a 208 μm thick $Cd_{0.23}Hg_{0.77}Te$ etalon at 300K for different initial detunings, B: +.25π, C: 0, D: -.2π, E: -.5π, F: -.75π. Peak incident intensity 150 kW/cm^2.
(b). Calculated transmitted pulse shapes for the same cavity conditions.

this is confirmed in the computation. These are due to the fact that the cavity takes a finite time to build up the required density of carriers to achieve the phase change necessary to establish the etalon in an upper state. This upper state is normally beyond the point of maximum transmission and during the build up period the cavity passes briefly through resonance before reaching the upper state. The speed of switching and overshoot behaviour depends on the rate of increase of intensity and the carrier recombination rate, as well as the cavity finesse and initial detuning. This is complicated in CdHgTe by the fact that the carrier lifetime is decreasing as the density of carriers is increasing, Eq. (5). The overshoots in the experimental results, Figure 6(a), are not so distinct which may be due to a combination of the limited detector response times, ~ 2 ns, not quite having plane wave illumination and inhomogeneities in the CdHgTe over the area sampled.

Figure 7(a) shows incident and transmitted pulse profiles for a single initial cavity detuning of a 363 μm sample using 1.75 μs pulses at a peak intensity of 500 kW/cm^2. A single switch up and subsequent recovery of the etalon is observed, B. Since the pulse length is far longer than the carrier recombination lifetime, the material is able to respond to the incident radiation in a quasi-cw fashion. Departure from exact cw behaviour is seen by the non-vertical switching; expanding the time base showed the rise to occur in ~ 50 ns. Figure 7(a) also shows the output from the computer program for the same experimental situation with the peak intensity scaled to 110 kW/cm^2. The simulation shows good agreement with measurements of the switch up but predicts a faster switch down than is observed. Figure 7(b) shows the cavity hysteresis loop (optical bistability) obtained from Figure 7(a) ensuring proper time matching of the input and output pulses.

5. BAND GAP DEPENDENCES

An important advantage of CdHgTe is that the band gap energy can be selected by an appropriate choice of alloy composition. In this section, we calculate the band gap dependencies of the optical nonlinearities at 10.6 μm for plane waves based on the agreement between experiment and theory for the compositions already studied. We consider the two cases of single photon resonance at 77K and two-photon excitation at 295K.

At 77K, we have to consider the dependence with band gap energy of (a) the absorption coefficient, (b) the carrier lifetime (and its dependence with excess carrier density, ΔN, Eq.(4)) and (c) the nonlinear refraction. Although it is unclear what physical mechanism gives rise to the band tail, Finkman and Nemirovsky[26] have found the absorption coefficient to be consistent with the expression,

the same functional variation with power, consistent with Auger recombination. The fact that the curves appear parallel indicates that the magnitude of the nonlinearity does not vary over the range of wavelengths used (10.51 to 10.69 μm). These results are consistent with a nonlinearity, $\gamma = -3.3 \times 10^{-3}$ $(cm^2/W)^{1/3}$ compared to a calculated value of -5×10^{-3}.

Figure 5 shows the reflected power as a function of incident power for the same sample after gold coating the back surface but retaining the uncoated front surface. A

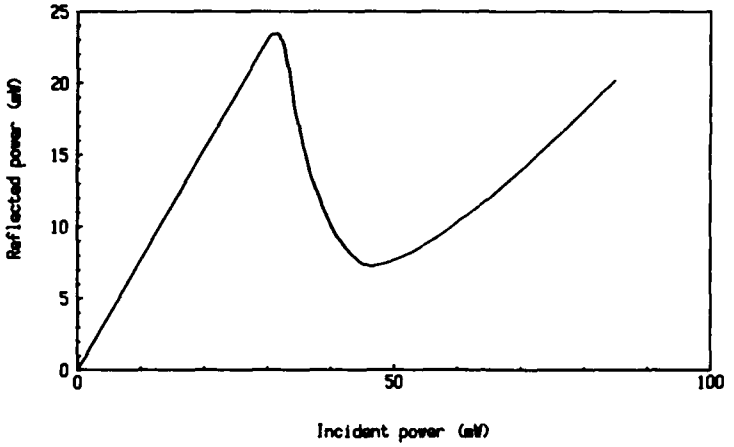


FIGURE 5. Input-output characteristic in reflection of 10.49 μm radiation with the use of a 208 μm thick $Cd_{0.23}Hg_{0.77}Te$ etalon at 77K with a high reflectivity coating on the rear surface.

20:1 mark-space chopper was placed in the incident beam to reduce thermal contributions. The increased finesse allows a strongly nonlinear behaviour which shows an optical gain of 2.2 at 30-40 mW. This characteristic points to the possibility of using CdHgTe in all-optical logic, modulation or differential amplification of CO_2 laser radiation.

Steeper characteristics were found to be difficult to attain due to the thermal properties of CdHgTe. This alloy is fairly unique for a zinc blende semiconductor in that heating increases the band gap energy resulting in increased transparency and, more importantly, a negative refractive index temperature coefficient. This is the same sign as the electronic nonlinearity. Thus, as a cavity approaches resonance due to electronic contributions to the refractive index at higher incident intensities, the power absorbed increases rapidly, raising the temperature of the material and so pushes the etalon further into resonance. The curve shown in Figure 5 is stable because the use of a 20:1 ratio chopper allows the absorbed power to dissipate during the period when the radiation is blocked. Operation of a nonlinear device in CdHgTe thus necessitates chopped or pulsed radiation.

4. TWO-PHOTON OPTICAL SWITCHING AND BISTABILITY

Experiments on two-photon induced nonlinear etalon effects in CdHgTe[8,9] were carried out at room temperature on crystals of the same composition, $x \sim 0.23$. This gives a band gap energy of 0.2 eV (6.2 μm). The two-photon absorption coefficient, β, in the material is relatively large. A theoretical analysis[21,22] results in a predicted value of $\beta \sim 4.8$ cm/mW. Using the band gap dependence of the two-photon absorption coefficient from the theory[21], extrapolation of measurements of β on a CdHgTe sample with a slightly different composition[23,24], gives $\beta \sim 6$ cm/mW.

Two separate lasers offering different pulse widths were employed. Both gave pulses with smooth temporal and spatial profiles. A short cavity, CO_2 TEA laser gave 30 ns (fwhm) duration pulses which are comparable in timescale with the intrinsic room temperature recombination lifetime of ~ 90 ns[17]. A hybrid CO_2 TEA laser gave 1.75 μs (fwhm) pulses which are longer than the recombination lifetime.

Figure 6(a) shows the changes in transmitted pulse shapes for different initial cavity detunings (obtained by sample rotation) using 30 ns pulses at a constant (unfocused) peak incident intensity of 150 kW/cm^2. Only the central part of the beam was sampled with a pinhole in front of the detector. Peaks in the profiles show that the cavity sweeps through two etalon transmission maxima. As the sample is rotated to larger angles, B to F, the nonlinear etalon features occur earlier in the pulse indicating the presence of a negative nonlinearity. The angle of rotation giving repetition of the pulse shape in B and F agrees with that expected for a 208 μm etalon.

Figure 6(b) shows results from a computer simulation[9] using the same cavity conditions as in Figure 6(a) other than the peak incident intensity which was adjusted to obtain a set of pulse profiles with features in agreement with the measurements. Because the recombination time of the excess carriers is of the same order as the pulse length, we would expect to observe transient switching phenomena such as transmission overshoots[25].

describe a semiconductor slice polished plane parallel with surface reflectivities, $R_F = R_B = 32\%$ and α chosen to be unity. Dimensionless intensity parameters are used in both cases and the initial cavity detuning has been adjusted to give the critical conditions for the lowest power optical bistability. It can be seen that two effects make optical bistability more difficult to obtain in the case of a semiconductor exhibiting Auger recombination, (i) the spacing of the cavity orders become larger and (ii) the slopes of the transmission curves become shallower with intensity in the higher orders. However, suitable optimised cavity design can significantly reduce the minimum incident intensity required for bistability[15].

Two-photon absorption adds another nonlinear process which additionally affects the power dependence of the nonlinearity. Under steady state conditions, the excited carrier density is given by

$$\frac{d\Delta n}{dt} = \frac{\beta I^2}{2\hbar\omega} - r\Delta N = 0 \qquad (14)$$

where β is the two-photon absorption coefficient. At room temperature, $Cd_{0.23}Hg_{0.77}Te$ is intrinsic and the recombination rate is given by

$$r = \left(1 + \frac{\Delta N}{N_i}\right)\left(2 + \frac{\Delta N}{N_i}\right) \Big/ 2\tau_i \qquad (15)$$

where $N_i = 1.5 \times 10^{16}$ cm^{-3} and $\tau_i \sim 90$ nsec. The refractive index should thus show a two thirds power dependence on intensity.

We can estimate that to decrease the optical thickness of a 200 μm thick etalon by $\lambda/2$ requires 2.3×10^{16} cm^{-3} generated carriers. This reduces the lifetime to ~ 20 nsec. The two-photon absorption coefficient, β, for $Cd_{0.23}Hg_{0.77}Te$ at 10.6 μm is ~ 5 cm/mW which requires a cavity intensity of ~ 100 kW/cm^2 to shift the etalon transmission by a full order.

3. SINGLE-PHOTON NONLINEAR FABRY-PEROT

Experiments were performed on an n-type (N_D-$N_A = 1 \times 10^{15}$ cm^{-3}) sample of $Cd_{0.23}Hg_{0.77}$ Te polished plane parallel to form an etalon 208 μm thick and held in a cryostat at 77K. The crystal absorption was 24 cm^{-1} in the 10 μm band with surface reflectivities of 32%. From Eq.(1), the band gap energy was estimated to be 0.13 eV. Radiation from a 2.5 W, line tunable CO_2 laser was passed through a variable attenuator-spatial filter arrangement to ensure a near Gaussian transverse profile and focused by a 20 cm focal length lens to a beam waist at the sample of diameter ~ 90 μm (1/e points of intensity profile). To avoid effects due to intensity dependent beam profile changes, the total transmitted or reflected laser power was measured using a collecting lens and a large area pyroelectric detector.

Figure 4(a) shows the highly nonlinear behaviour of the crystal using a low power cw beam. Rapid transmission changes can be observed at power levels as low as 100 μW. In

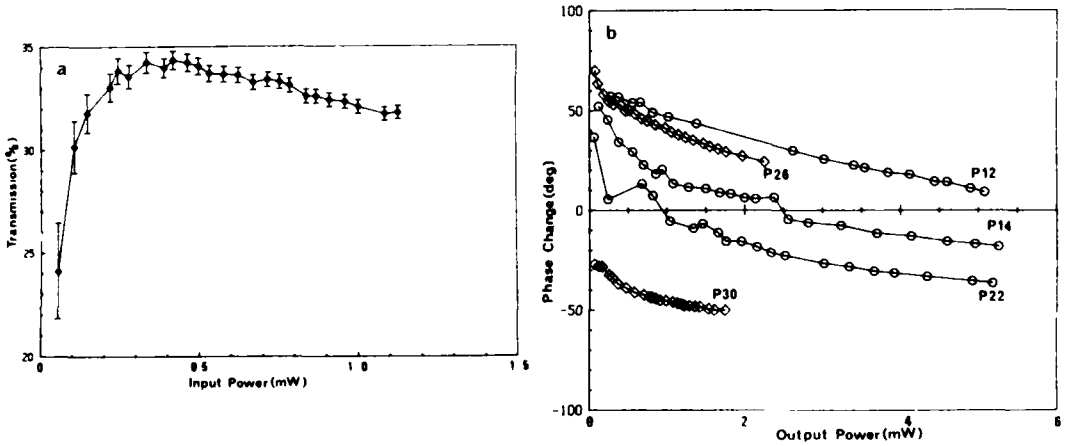


FIGURE 4(a). Transmission of cw 10.6 μm radiation through a 208 μm thick $Cd_{0.23}Hg_{0.77}Te$ etalon at 77K.
(b). Plot of intensity dependent phase changes induced in the etalon at five different wavelengths of the CO_2 laser.

the case shown, transmission starts at a low value and increases with intensity; at other wavelengths where the cavity is initially on resonance, the transmission decreases first. The transmission can be seen to vary more rapidly at the lowest powers consistent with the expected 1/3 power behaviour. From transmission versus input power curves at different wavelengths, the variations of induced cavity phase change with output power (which is proportional to the cavity intensity) are plotted in Figure 4(b). The data measured show

The transmission, T, and reflection, R, of a Fabry-Perot of length, ℓ, with front and back reflectivities, R_F, R_B, containing material of linear absorption, α, is given by,

$$T = \frac{A}{1 + F \sin^2 \phi} \tag{10}$$

$$R = \frac{B + F \sin^2 \phi}{1 + F \sin^2 \phi} \tag{11}$$

where

$$F = \frac{4R_s}{(1-R_s)^2}$$

$$A = \frac{e^{-\alpha \ell}(1-R_F)(1-R_B)}{(1-R_s)^2}$$

$$B = \frac{R_F(1 - (R_s/R_F))^2}{(1-R_s)^2}$$

and

$$R_s = (R_F R_B)^{\frac{1}{2}} e^{-\alpha \ell} \quad .$$

The intensity dependence of transmission and reflection comes about through the phase, ϕ, which can be divided into an initial detuning of the cavity, ϕ_0, and an intensity dependent part,

$$\phi = \frac{2\pi n \ell}{\lambda} = \phi_0 + \frac{2\pi \Delta n \ell}{\lambda} \quad . \tag{12}$$

With an absorbing material, the cavity transmission (or reflection) and input intensity, I_i, can be shown[20] to be related through the effective mean internal cavity intensity, I_c,

$$I_i = C_T \frac{I_c}{T} = C_R \frac{I_c}{(1-R)} \tag{13}$$

where

$$C_T = \frac{\alpha \ell e^{-\alpha \ell}(1-R_B)}{(1-e^{-\alpha \ell})(1+R_B e^{-\alpha \ell})} \quad \text{and} \quad C_R = \frac{\alpha \ell \left(1 - \frac{R_s^2}{R_F}\right)}{(1-e^{-\alpha \ell})(1+R_B e^{-\alpha \ell})} \quad .$$

Eqs.(10) and (13) solved simultaneously to eliminate I_c describe nonlinear Fabry-Perot transmission as a function of incident intensity. The existence of optical bistability is normally determined graphically by observing whether a multiple intersection occurs between the periodic dependence of T on I_c from Eq.(10) and the straight line defined by Eq.(13).

Figure 3 shows plots of Eqs.(10) and (13) for a semiconductor with (a) density independent recombination rate, $\Delta n \propto I$ and (b) Auger recombination, $\Delta n \propto I^{1/3}$. Both figures

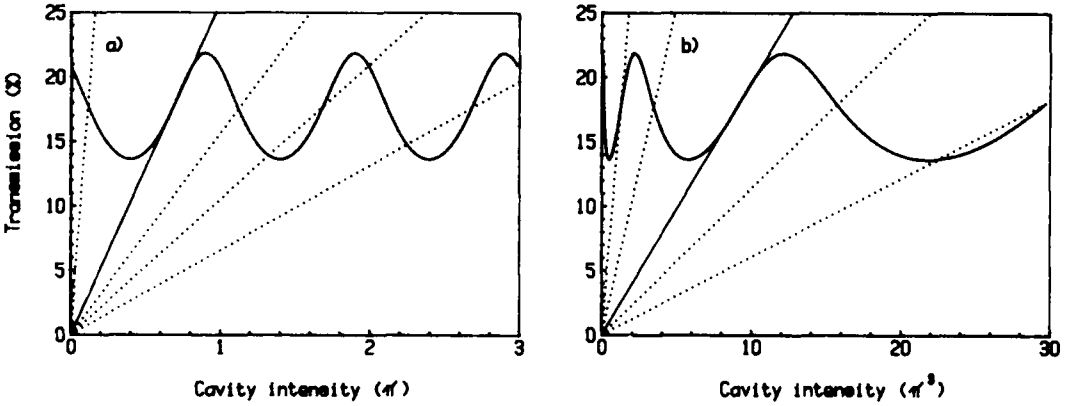


FIGURE 3. Nonlinear Fabry-Perot transmission versus mean cavity intensity for $R_F = R_B = 32\%$ and $\alpha \ell = 1$ for (a) $\Delta n \propto I$ and (b) $\Delta n \propto I^{1/3}$.

where I is the intensity in the material and α is the absorption coefficient. A reasonable approximation of Eq.(4) for $\Delta N > N_0$ gives the intensity dependence of the excess carriers,

$$\Delta N = \left[\frac{2N_1^3 \alpha I}{\hbar\omega}\right]^{1/3} \tag{6}$$

and from Eq.(3), the refractive index change has the form,

$$\Delta n = \gamma I^{1/3} \quad . \tag{7}$$

This intensity dependence was confirmed in self-defocusing studies of 10.6 μm radiation in a 330 μm thick sample of $Cd_{0.21}Hg_{0.79}Te$ at 175K[5]. The one third law manifests itself in the power dependence of the far field beam profile as shown in Figure 2. For a

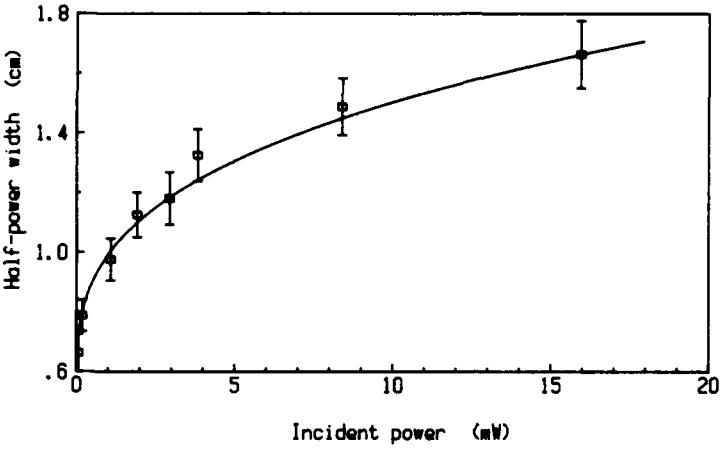


FIGURE 2. Power dependence of far-field beamwidths for 10.6 μm radiation after transmission through 330 μm thick $Cd_{0.21}Hg_{0.79}Te$ at 175K placed just beyond a beam waist of 200 μm (fwhm). The solid line shows a one-third power dependence for comparison.

Gaussian beam profile and a uniform absorption coefficient, the phase change across the beam profile due to the nonlinear interaction can be calculated from,

$$\Delta\phi(r) = \frac{2\pi}{\lambda} \int_0^{\ell} \gamma \Delta N(r,z)\, dz$$

and in this case, from Eqs.(3) and (7),

$$\Delta\phi(r) = \frac{2\pi}{\lambda} \frac{3\gamma(1-\exp(-\alpha\ell/3))}{\alpha} I_o^{1/3} \exp\left(-\frac{r^2}{3d^2}\right) \tag{8}$$

where I_0 is the central beam intensity just inside the sample, d is the width parameter of the incident beam, $I_0 \exp(-r^2/d^2)$, and ℓ is the sample thickness. The profile of phase variation is $\sqrt{3}$ times broader than the intensity profile. This allows the phase variation to be expanded as a power sum of quadratic variation over the most significant part of the intensity profile. Retention of only the leading quadratic term leads conveniently to a representation in which the phase angle variation approximates to that of a concave lens of focal length,

$$\frac{1}{f} = \frac{2\gamma(1-\exp(-\alpha L/3))}{\alpha} \frac{I_o^{1/3}}{d^2} \quad . \tag{9}$$

The far field profile then remains Gaussian but its width is determined by both the intensity dependent interaction in the CdHgTe and the propagation of the beam prior to the sample.

The retention of a Gaussian beam profile on transmission is advantageous to the operation of optically bistable devices and may be compared with materials with n$_2$I-type nonlinearities in which considerable beam break-up appears[18,19]. However, the one third power dependence makes the conditions for optical bistability more difficult to attain.

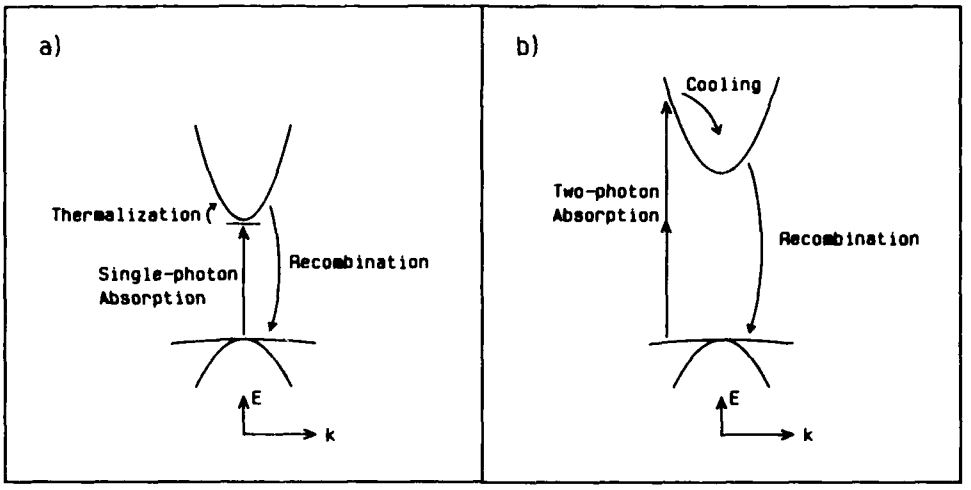


FIGURE 1. Excitation mechanisms for dynamic nonlinear optical effects at 10.6 μm in $Cd_{0.23}Hg_{0.77}Te$, (a) 77K, (b) 295K

volume, γ, then,

$$\gamma = -\frac{e^2}{2\varepsilon_o m^* n \lambda^2}\left[1 + \frac{2}{3\pi}\frac{m^*}{m}\left(\frac{mP^2}{\hbar^2}\right)\frac{1}{kT}J\left(\frac{\hbar\omega - E_g}{kT}\right)\right] \qquad (2)$$

where

$$J(a) = \int_0^\infty dx\, x^{\frac{1}{2}}\, e^{-x}/(x-a)$$

and

$$\Delta n = \gamma \Delta N \qquad . \qquad (3)$$

Here, e is the electronic charge, ε_o the free space dielectric constant, m* the conduction band effective mass, m is the electron mass, $\hbar\omega$ is the photon energy, n the linear refractive index and P the momentum matrix element (the quantity, $mP^2/\hbar^2$ is in Joules). The first term describes the plasma contribution to the nonlinear refraction and is the Drude expression derived from standard dispersion theory modified by the electron effective mass[13]. The second term is the result of the excess carriers blocking virtual transitions between the valence and conduction bands[14] and shows a resonance for photon energies close to the band gap, embodied in the thermodynamic integral, J(a).

The advantage of employing long wavelengths and small gap semiconductors are apparent in Eq.(2) from the $1/\lambda^2$ dependence. The momentum matrix element is fairly constant for different zinc blende materials while the electron effective mass is approximately proportional to the band gap energy. For the 77K, $Cd_{0.23}Hg_{0.77}Te$ samples employed in the experiments described in section 3 under 10.6 μm band gap resonant excitation, we would expect, $\gamma \approx -2.3 \times 10^{-18}$ cm^3. The same composition at room temperature with E_g = 0.2 eV should give $\gamma \approx -1 \times 10^{-18}$ cm^3 at 10.6 μm.

The density of carriers in a semiconductor under optical excitation is governed by the dynamic balance of generation, recombination and diffusion. Under the conditions employed in our experiments, diffusion can be neglected. Recombination in CdHgTe at both 77 and 295K is dominated by Auger processes[4] which cause a strong concentration-dependence for the carrier lifetime. This therefore has a significant effect on the intensity dependence of the nonlinearity[5] and in turn the form of the nonlinear Fabry-Perot characteristics[15].

For our 77K sample, with an initial electron doping, $N_o = 1 \times 10^{15}$ cm^{-3}, the recombination rate is given by[16],

$$r = \frac{\Delta N(N_o + \Delta N)^2}{2N_i^2 \tau_i} \qquad (4)$$

where N_i and τ_i are the intrinsic carrier concentration and lifetime respectively. In this case, $N_i^2\tau_i \sim 1.3 \times 10^{24}$ cm^{-6} s which gives a low intensity carrier lifetime ($\Delta N \ll N_o$). of ~ 2.5 μs[17]. The density of excited electron-hole pairs under steady-state conditions is given by,

$$\frac{d\,\Delta N}{dt} = \frac{\alpha I}{\hbar\omega} - r\Delta N = 0 \qquad (5)$$

OPTICAL BISTABILITY IN $Cd_xHg_{1-x}Te$

A. Miller, D. Craig, G. Parry*

Royal Signals and Radar Establishment,
Malvern, WR14 3PS, UK

J.G.H. Mathew, A.K. Kar

Department of Physics,
Heriot-Watt University,
Edinburgh, EH14 4AS, UK

SUMMARY

We discuss the origin and characteristics of nonlinear effects in CdHgTe Fabry-Perot etalons at 10.6 μm which arise from large intensity dependent refractive index changes. Optical switching and bistability can be observed in room temperature samples by two photon excitation. The band gap dependences of these phenomena are analysed.

1. INTRODUCTION

Semiconductor materials show great promise for all-optical signal processing and computing applications. The extremely large, band gap resonant, nonlinear optical effects discovered in recent years[1] have allowed the demonstration of low power optical bistability and related phenomena in a number of compound semiconductors[2,3]. The electronic interactions can be very fast and combined with the parallelism that optics can offer, semiconductors may well provide ultrafast signal processing elements.

These semiconductor optical nonlinearities normally take the form of refractive index changes caused by excitation of electrons across the fundamental band gap. Theory predicts that these effects should be larger for small band gap materials using correspondingly longer wavelength radiation. The alloy semiconductor, $Cd_xHg_{1-x}Te$ offers a variation of band gap energy with composition and temperature from 0 to 1.5 eV[4]. This wide range thus offers energy gaps resonant with a number of important laser frequencies and in particular it is one of the few semiconductors suitable for the study of band gap resonant optical nonlinearities in the important CO_2 laser output band around 10 μm. Indeed, we have observed very large nonlinear refractive effects of electronic origin in CdHgTe under resonant conditions using ~ 1 mW from a cw CO_2 laser[5]. In recent publications, we have reported nonlinear Fabry-Perot transmission and reflectivity at mW power levels in a cooled crystal[6,7] as well as optical switching and bistability at room temperature using two-photon excitation with pulsed lasers[8,9]. In this paper, we review these optical nonlinearities in Fabry-Perot etalons of CdHgTe and consider the band gap dependences of these effects for 10.6 μm radiation.

2. NONLINEAR REFRACTION IN CdHgTe

Moderate densities of optically generated electrons and holes can cause relatively large changes in the refractive index of small gap semiconductors. Figure 1 illustrates two methods of generating carriers which we have studied both experimentally and theoretically in CdHgTe. In (a), single photon absorption excites carriers into bandtail states. These carriers are scattered into the band within a few picoseconds and will recombine on longer timescales, usually within nanoseconds or microseconds. In (b), two photon absorption excites electrons directly into the conduction band with some excess kinetic energy. These carriers drop to the bottom of the band by phonon emission on a picosecond timescale before recombining. Although the two photon excitation process is not as efficient as single photon absorption for the generation of carriers, smaller gap semiconductors have relatively large two photon absorption coefficients. In addition the excited carriers are normally more evenly distributed through the material and the background absorption can be much smaller than for single photon excitation. CdHgTe is unusual for a zinc blende semiconductor in that the band gap increases in energy with increasing temperature. The band gap dependence on composition, x, and temperature, T, for $Cd_xHg_{1-x}Te$ has recently been found by Hansen et al[10] to be consistent with the expression,

$$E_g(x,T) = -0.302 + 1.93x + 5.35(10^{-4})T(1-2x) - 0.810x^2 + 0.832\,x^3 \qquad (1)$$

Thus, a sample of alloy composition which has a low temperature band gap matched to the output of the CO_2 laser at 10.6 μm, $x \simeq 0.23$, will be suitable for two photon excitation at room temperature.

The refractive index change, Δn, for photon energies below the band gap energy, due to an excess free carrier density, ΔN, can be considered to arise from two contributions[11,12]. Expressing the nonlinearity as a change in refractive index per electron-hole pair per unit

---

*Present address: Department of Electronic and Electrical Engineering, University College, London WC1E 7JE, UK.

Results

Optical Modulators - Bifurcation to Chaos in NH_3

Ikeda oscillation has been observed by use of a number of laser transitions nearly coincident with NH_3 absorption lines. Here we concentrate on the 10R(14) (CO_2 line [illegible] = 10.3 μm), pumping the aR(1,1) NH_3 transition 1.2[illegible] GHz above line centre. Over most of the pressure range (1.5 - 40 torr) the system acts as an effectively pumped, homogeneously broadened 2-level system, the pressure broadened full width at half maximum (FWHM) being 2.7 MHz $torr^{-1}$.

The transmitted signal showed large pulse distortion and modulation, the structure of which was sensitive to NH_3 gas pressure, cavity tuning, and input-signal intensity. The modulation period of the transmitted signal scaled linearly with cavity length, as expected.

Representative examples of the modulated output for PZT tuning are shown in Fig. 3(b) for a cavity length 86 cm, cell length 70 cm, and pressure 10 torr, with input-coupler reflectivity 64%. Here no Ikeda oscillation (period ~ 13 ns, close to $2t_R$ = 11.5 ns) persistent throughout the pulse is evident in the neighbourhood of minimum transmission, consistent with the above mentioned four-wave mixing interpretation of this instability (9). PZT tuning of the cavity leads progressively to "switching" behaviour with high peak transmission followed by damped oscillation of longer period. The corresponding input shown in Fig. 3(a) is for the same cavity without gas.

At lower pressures (4 - 8 torr), where inhomogeneous broadening may be important, strong $4t_R$ oscillation was observed. At higher pressures (20 - 35 torr) much more complex pulse shapes were obtained. These features were enhanced for reduced input-coupler reflectivity (R_o = 36%); large input coupling being needed to bleach the high absorption ($\alpha L \sim 6$) to achieve adequate cavity feedback. PZT sequences for pressures ~ 20 torr showed $2t_R$ oscillation, developing to $2t_R/3$ oscillation on the higher branch, and bifurcating to $4t_R/3$ before again evolving to lower-branch $2t_R$ modulation. Aperiodic pulse shapes, characteristic of chaos, were also evident. Similar behaviour was obtained on a number of other transitions in NH_3 *in close resonance with* CO_2 *laser lines.*

It is significant that the $2t_R$ oscillation was the most reproducible and rather insensitive to parameter control. This is consistent with predictions (11) and is especially encouraging in establishing such systems as passive, high frequency modulators, tunable by cavity length control.

Stability analysis of nonlinear Fabry-Perot resonators has apparently only been examined in the dispersive (Kerr) limit (12). Our experiments, however, necessarily involve the bleaching of a rather substantial absorption. We have therefore adapted the model of Carmichael & Hermann (13) which treats steady-state optical bistability in a Fabry-Perot resonator. We handle the time dependence by replacing the gas cell by N thin slices symmetrically placed within the cavity. Assuming the adiabatic limit, we apply steady-state theory to find the forward and backward transmissions of each slice (unequal due to the phase-population grating). We can then follow the field around the cavity, where it interacts in the slices with N earlier and N later fields, and keep track of its attenuation at each stage. At the output and input couplers we apply the usual Fabry-Perot boundary conditions, and thus have, in effect, a 2N + 1 parameter mapping problem. Full details of this analysis are reported elsewhere (7).

Application of this procedure to NH_3 and use of an input pulse of the form shown in Fig. 2 yields the transmitted pulse shapes [illegible] ig. 3(c) in pleasing agreement with the observed pulse shapes, especially since only measured parameters are used: α is 0.025 cm^{-1} at 10 torr and scales as p^2 (this, plus the pressure broadening rate and constants gives a saturation intensity I_s ~ 2.3 MW cm^{-2} at 10 torr (1.5 MW cm^{-2} at 40 torr)). The value I/I_s ~ 7 is thus in line with the measured input intensities in the range 10 - 20 MW cm^{-2}.

This good agreement is rather surprising in view of the omission of reservoir (6,7) and transverse effects in the analysis. Inclusion of such effects, in particular transverse (14) is, however, necessary to generate the more complex pulse shapes obtained at higher pressure, though qualitative agreement still exists from the plane-wave approach.

Compact Systems - Optical Hysteresis in SF_6

It is emphasised that for cw inputs these models give steady-state oscillation and chaos within our experimental parameter range and we have identified period-doubling to $16t_R$ en route to chaos in one case. An example of $2t_R$ oscillation in NH_3 for an input power of ~ 6 I_S is shown in Fig. 4. The corresponding power spectrum also shows contributions at t_R and $2t_R/3$; a manifestation of the non-sinusoidal form of the oscillation. In practice, generation of these effects under cw conditions prescribes the use of molecules with low saturation intensities, such as SF_6 (ca. 6 W cm^{-2} $torr^{-1}$). As a precursor to this we report observation of optical hysteresis effects in this gas over a wide range of parameter conditions. We note that operation here is in the 'bad cavity' limit and so precludes generation of Ikeda instabilities. Fabry-Perot resonators, ranging in length from 2 to 15 cm, with intra-cavity gas cells as short as 1 mm were operated at SF_6 gas pressures ranging from ca. 1 to ca. 200 torr. Various lines in the 10P band of the TEA CO_2 laser were used to pump the dense and broad spectral features of the ν_3 vibrational mode of SF_6 near resonance. Effects of switching, power limiting, and overshoot with nanosecond response times, limited only by cavity decay time, were routinely obtained at gas pressures commensurate with self focussing. A sample trace of the transmitted signal is shown in Fig. 5(b) for an SF_6 pressure of ca. 10 torr pumped at 947.7 cm^{-1} (10P16 CO_2 line). Corresponding recordings of instantaneous input (x axis) and output (y axis) intensity (Fig. 5(c)) show good bistable action. Equivalent recordings for the empty cavity are shown in Fig. 5(a) for reference.

The successful miniaturisation of systems here is again encouraging with regard to device applications as are our recent observations of saturation in SF_6 using a 3 W cw input signal. We note generation of Ikeda modulation in such systems will, however, require higher operating pressures to ensure a relaxation

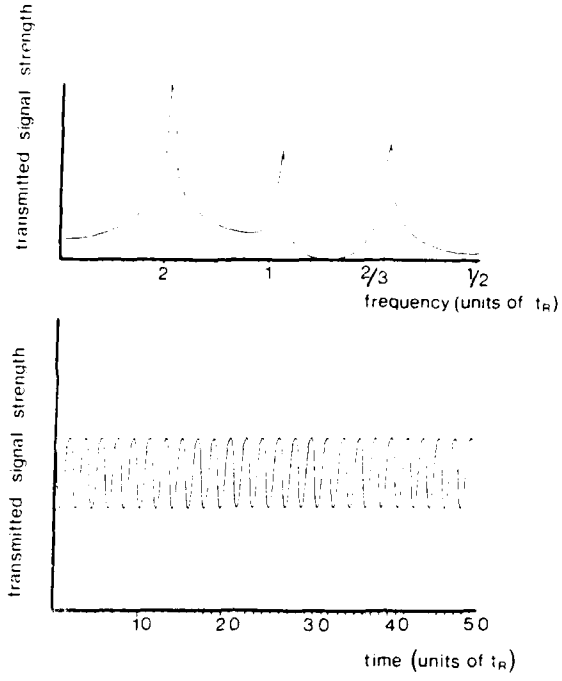


Fig. 4. Computer traces of the transmitted signal (lower) and power spectrum (upper) for cw input signal; power 6.5 I_s for NH_3 pressure of 20 torr and cavity detuning of 3.5 radians.

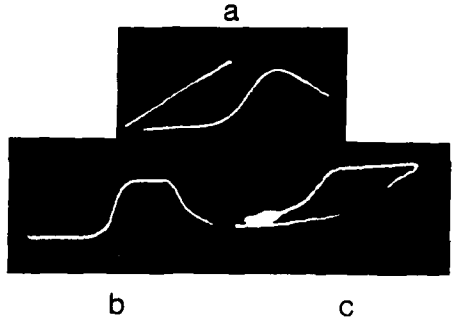


Fig. 5. Oscilloscope recording of transmitted signal together with resultant trace for simultaneous recording of input and transmitted signals for a) empty cavity and b) and c) cavity with SF_6 at pressure 5.5 torr cell length 10 cm pumped at 10.6 μm.

time of the gas shorter than the round trip time of the cavity. Here the use of buffer gas is envisaged.

In conclusion, our observations of $2t_R$ oscillation period-doubling to $4t_R$ and, over a more limited parameter range, $2/3t_R$, $4/4t_R$ modulation and aperiodic pulse shapes in NH_3 gas confirm the predicted behaviour of this system and support the conclusion that we have driven an all optical 2-level system through oscillation to chaos. Already of use as a passive optical modulator ($2t_R \simeq 100$ MHz) of pulsed signals, prospects for modulation in the gigahertz range are promising. The possibility of generating these effects under cw conditions in compact systems is supported by our observations of strong optical hysteresis in this system by using SF_6 which, along with many other molecular gases, exhibits low saturation intensity.

This work is supported by the Science and Engineering Research Council, and in part by the European Joint Optical Bistability programme. I.A.A. acknowledges assistance from the Government of Iraq.

References

(1) D.A.B. Miller and S.D. Smith, Opt. Commun., 31, 101 (1979).

(2) K. Ikeda, Opt. Commun., 30, 257 (1979).

(3) H. Nakatsuka, S. Asaka, M. Itoh, K. Ikeda and M. Matsuoka, Phys. Rev. Lett., 50, 109 (1983).

(4) H.M. Gibbs, F.A. Hopf, D.L. Kaplan and R.L. Shoemaker, Phys. Rev. Lett., 46, 474 (1981).

(5) R.G. Harrison, W.J. Firth, C.A. Emshary and I.A. Al-Saidi, Phys. Rev. Lett., 51, 562 (1983).

(6) R.G. Harrison, W.J. Firth, C.A. Emshary and I.A. Al-Saidi, Appl. Phys. Lett., 44, 718 (1984).

(7) R.G. Harrison, W.J. Firth and I.A. Al-Saidi, Phys. Rev. Lett., 53, 258 (1984).

(8) R.G. Harrison, I.A. Al-Saidi and W.J. Firth, Phil. Trans. R. Soc. Lond. (in press).
R.G. Harrison, I.A. Al-Saidi, E.J.D. Cummins and W.J. Firth, Appl. Phys. Lett. (submitted).

(9) W.J. Firth, E.M. Wright and E.J.D. Cummins, Optical Bistability 2, eds. C.M. Bowden, H.M. Gibbs and S.L. McCall, Plenum Press, pp. 111-118 (1984).
W.J. Firth, E.M. Wright and E. Abraham, JOSA, B1, 489-90 (1984).

(10) K. Ikeda, H. Daido and O. Akimoto, Phys. Rev. Lett., 45, 709 (1980).

(11) M.J. Feigenbaum, J. Stat. Phys., 19, 25 (1978).

(12) W.J. Firth, Opt. Commun., 39, 343 (1981).
E. Abraham, W.J. Firth and J. Carr, Phys. Lett., 91A, 47 (1982).

(13) H.J. Carmichael and J.A. Herman, Z. Phys., B38, 365 (1980).

(14) J.V. Moloney, F.A. Hopf and H.M. Gibbs, Phys. Rev. A, 25, 3442 (1982).

DISCUSSION

M.Weiss, US

I wonder whether you've tried any asymmetric molecule like partially deuterated ammonia (NH_2D or ND_2H) which should permit you to use it at a variety of wavelengths?

Author's Reply

Even with ordinary ammonia NH_3 we've actually achieved similar results on probably four or five transitions. The incentive of going to deuterated ammonia is probably not there on that basis. We've found that the results I've shown here are in fact quite general for off-resonance excitation of transitions, not only in ammonia but other molecules as well.

J.W.Haus, US

This chaotic output you mentioned as you tuned through the frequency — it looked intermittent to me, not really chaotic. There were some areas where it looked nearly the same, then it would glitch a bit, then look the same again for a while. That's also expected in the Ikeda theory. Could it have jumped over the chaotic region and into the intermittent?

Author's Reply

It's difficult. Because these are actually pulse experiments one is actually taking Ikeda's analysis and applying a dynamic situation to it anyway. And as you probably know it is quite difficult to unravel the basic signs. Computer analysis is involved in solving the Ikeda equations so in unravelling the basic structures at a quantitative level is quite tricky.

J.W.Haus, US

We analysed the Ikeda equations for various cavity detunings in this ring cavity. We found that we had a coexistence of steady-state lower state with a period-2 upper state in bistability between them. That was possible with an appropriate cavity detuning. Do you think you'd be able to distinguish anything like that?

Author's Reply

Pulse experiments — again, probably not. Even in some of the traces I showed today you tend to get a switching effect where the oscillation sometimes occurs on the upper branch, although in others at lower pressures you can actually get a small degree of oscillation on the lower branch; then it switches. There's a difficulty in unravelling again qualitative features of this type in pulse experiments. In a cw experiment your control parameter could be your laser intensity so you'd actually operate on the lower branch, look for instabilities and then increase the intensity to switch up and then hope to see the instabilities again. This is happening automatically in a pulse experiment where you have a ramp in intensity of your input signal and then a drop. So you have this difficulty.

J.W.Haus, US

What about transverse effects? Moloney has made a lot of calculations.

Author's Reply

Yes. We haven't as yet seen or ratified those predictions, but then we haven't been sampling with an aperture the parts of the spatial cross-section of the output beam. We've been collecting the whole signal. Moloney has been concentrating on various regions of the cross-section of the output signal.

J.E.Carroll, UK

What is your theoretical switching speed and what is the relationship of that speed to the question of observing chaos? The point I'm making here is that it was too slow — I doubt that you could observe chaos.

Author's Reply

Our switching times in SF_6 are of the order of nanoseconds or so which in fact were longer than the round trip of the cavity, and therefore we did not see bifurcation to chaotic emission. In ammonia the response times of the medium were the order of nanoseconds and possibly slightly less under all pressure conditions and thus considerably less than the round-trip time of the cavity. The switching speeds would have therefore been faster because they're essentially determined by the cavity finesse and the input power of the laser signal.

J.E.Carroll, UK

Couldn't this have been the problem in your final experiment where you didn't observe chaos in your much shorter cavity?

Author's Reply

Indeed. That's what a bad-cavity system is — that the response time of your medium is, in fact, actually longer than the round trip of your cavity.

S.D.Smith, UK

You extrapolated the frequencies from the observed roundabout nanoseconds times up to 15 GHz. If one were to take it a bit further and question whether you could reach 100 GHz the practical application of it would be, of course, that that frequency could be controlled simply by controlling cavity lengths. That seems to be very good because if it could make a clock, it could drive computers and processors. As Andy Walker will show later, we've already seen MHz modulation with gain; that can also be extrapolated. One could speculate whether or not one could reach 100 GHz, because if you could it could get beyond I think the electrooptic modulator frequencies. And then it really would become interesting for frequency mixing type experiments, and devices, of course.

Authors's Reply

Well, we're already speculating to say we might be able to get to 15 GHz let alone higher. It's difficult because one's got to understand the molecular physics here. We think we understand it. To do it you'd have to go to very high pressure multiatmosphere pressure, in which case the absorbing molecules would make up a very small part of the total pressure of the system. You'd have to add a buffer gas to it. The problem you're faced with is the uncertainty of what's going to be the dependence on pressure of the saturation intensity. It is going to require extremely high input powers to achieve the necessary nonlinearity. What I might just point out is that the mechanism we're talking about is a vibrational rotational transition in the molecule. So that's a ground vibrational level, that's a first excited vibrational level — that box and that box are the rotational levels which are being coupled by the radiation off-resonately. There is nevertheless a strong interaction of population between the pump transition, the optically excited transition, and the reservoir of population in the other remaining rotational states. The effects we've been talking about are dictated by relaxation times associated with this transition here and here. The transfer of population from the bath to the pump level. That is a rotational relaxation process. If we're trying to go cw we must have a complete recirculation, and then we're dictated by the slowest rate constant which is a VT(?) process which is the order of microseconds. In order to make *that* extremely short you have to go to extremely high pressures. Now this opens up the question that we still have to answer: what happens to effects of saturation intensity. And even so, the arguments I've given here are quite simplistic because they're still specific to 2-level systems which are terrific in quantifying, again as I say, the Ikeda analysis but at a device level you're not probably going to go for those kind of systems. You're going to go for far more complicated molecules which exhibit very strong absorption cross-sections — large α_0's — which exhibit low saturation intensities. SF_6 is just a typical example of many — VCl_3, SiF_4 — there are stacks of them which are far superior in these criteria in comparison with ammonia. There's a lot of work to be done on that to determine it. But our results at this stage are positive in confirming that modulators can be made, and in fact the system that we have at the moment is already useful as a device system.

AN EXPERIMENTAL NONLINEAR OPTICAL WAVEGUIDE DEVICE

I Bennion, M J Goodwin, D J Robbins and W J Stewart

Plessey Research (Caswell) Limited
Allen Clark Research Centre
Caswell Towcester Northants NN12 8EQ

SUMMARY

An experimental nonlinear optical waveguide device is described which comprises a glass optical waveguide surmounted by CS_2. Intensity-dependent hysteresis in the power guided by the structure has been observed which is related to the intensity-dependent refractive index of the overlayer.

1. INTRODUCTION

It has been shown by Smith et al[1] that, under certain conditions of excitation, the interface between an optically linear and a nonlinear medium demonstrates reflection characteristics that exhibit hysteresis as a function of the incident light intensity. Interest in the nonlinear interface centres on its potential application in very fast optical switching and logic elements, for which it is hoped to utilise the subpicosecond response times of optical material nonlinearities. In the experiments reported here, the surface of a planar optical waveguide is contacted to a nonlinear medium to produce a nonlinear interface in a simple waveguide format. It is anticipated that a waveguide configuration, being compact and capable of achieving the high optical intensities required for nonlinear optics from a source of relatively modest power output, might form the basis for a device employing this effect.

We have analysed the waveguide nonlinear interface using a plane-wave approach similar to that employed by Kaplan[2] for the bulk case. The analysis shows that for the TE modes of the waveguide a straightforward generalisation of linear slab waveguide theory leads to modes parameterised by the power flow in the waveguide in the nonlinear case. This approach has some advantages over the alternative coupled-mode analysis; in particular, it may be applied to investigate changes in the mode structure near cutoff. The theory shows that the bistability predicted by Kaplan emerges as a bistability in the propagation constant of the waveguide mode, the two states lying in the guided and radiation mode regimes, respectively. The power flow assumes the role of control parameter in this model. We have not undertaken analysis of the stability of the two states, although a more recent experimental study of the bulk nonlinear interface suggests that one of the two states observed is unstable with a lifetime of the order 10^3 times longer than the response time of the material nonlinearity[3].

The experiment reported here is the first stage in a study of the nonlinear switching behaviour of this waveguide structure. Ultimately, it is desirable to utilise waves propagating strip or channel waveguides wherein the optical energy is confined within a cross-section of a few μm^2 compatible with high packing density and low power operation. For simplicity, however, the waveguide utilised here is of planar geometry, with light confined in only one dimension transverse to the direction of propagation.

In this paper, we summarise the plane-wave theory of the waveguide device in section 2, and describe the experimental realisation in sections 3 and 4.

2. PLANE-WAVE ANALYSIS OF THE NONLINEAR WAVEGUIDE

An analysis of TE mode propagation in an infinite slab waveguide bounded by a nonlinear medium has been given in ref.(4). A generalisation of established slab waveguide theory is described which is the evanescent wave solution of the nonlinear scalar wave equation given by Kaplan[2] in his treatment of reflection at a bulk nonlinear interface.

The structure under consideration is illustrated in Fig. 1. The optically linear guiding layer of refractive index n_1 and width $2a$ is supported on a substrate of refractive index n_2 and overlaid by a nonlinear medium of refractive index n_3. The nonlinear index n_3 is intensity dependent and described by

$$n_3(|E|^2) = n_{3l} + (2n_{3l}\Delta - \Delta^2) + \varepsilon_{n_3}|E|^2 \qquad (1)$$

where E is the electric field strength of the optical wave, and ε_{n_3} the positive optical Kerr coefficient.

We consider z-directed propagation of a single TE mode (E-field parallel to the waveguide) with

$$E_y = U(x) \cdot e^{-i(\omega t - \beta z)} \qquad (2)$$

where β is the (longitudinal) propagation constant of the wave. Solution of the scalar wave equation in the three regions of the structure and appropriate field matching at the boundaries leads to the generalised mode equation[4] :

$$\tan 2a\sqrt{k^2 - b} = \frac{\sqrt{k^2 - b}\,\left(\sqrt{\alpha + b - 1} + \sqrt{\beta + b - 0}\right)}{k^2 - b - \sqrt{\alpha + b - 1}\,\sqrt{\beta + b - 0}} \qquad (3)$$

where

$$\left.\begin{aligned} v^2 &= a^2k^2(2n_1\Delta - \Delta^2) \\ d &= \frac{a^2}{v^2}(n_1^2 - n_2^2)k^2 \\ b &= \frac{a^2}{v^2}(\beta^2 - n_1^2k^2 + (2n_1\Delta - \Delta^2)k^2) \end{aligned}\right\} \quad (4)$$

and $k = 2\pi/\lambda$, with λ the free-space wavelength. d is a measure of the asymmetry of the waveguide structure: for a symmetrical ($n_2 = n_3$), linear structure, $d = 1$. The intensity-dependence of the refractive index n_3, enters the mode equation (3) only through the parameter Q, which is given by

$$Q = \tfrac{1}{4}a^2k^2\epsilon_{nl}|E(0)|^2 \quad (5)$$

$|E(0)|^2$ is the intensity at the nonlinear interface, upon which both modal field distribution and power flow depend. When $Q = 0$, equation (3) reduces to the well-known linear waveguide result.

In the presence of absorption, the propagation constant becomes position-dependent, as discussed in ref.(5). We restrict the present analysis to cases where the nonlinear medium is lossless in order to preserve the validity of the evanescent field solution in that medium. In the experiments, this requirement is well fulfilled by the nonlinear material used, CS_2, at a wavelength of 1.06 μm.

A range of waveguides are considered in the present work and these lie in the asymmetry range $d = (5,10)$ typically, with $v \leq 1$. We illustrate the power dependence of the normalised propagation constant, b, in Fig. 2 for $d = 10$, $v = 1$. For a CS_2 overlayer, $P_{norm} \sim 4 \cdot 10^7$ W.cm^{-2}: this is obtained by solving the mode equation (3) parameterised by Q and then calculating the associated power flow.

Qualitatively, it may be seen that with increasing power flow, the refractive index in the nonlinear medium is progressively increased to a value at which the waveguide ceases to confine the propagating mode. The effect on the modal field distribution is shown in Fig. 3 for the same waveguide parameters $d = 10$, $v = 1$, as in Fig. 2. From equation (3) it follows that this cut-off point occurs when $v^2b = Q$ in the nonlinear case with guided modes restricted to the range $v^2b > Q$. The wave equation also admits plane wave solutions corresponding to radiation modes of the waveguide and these exist so long as $bv^2 < 2Q$. Provided that Q is non-zero, the two regions overlap. At the bulk nonlinear interface, it is the reflectivity that exhibits hysteresis with incident intensity as control parameter and angle of incidence as an order parameter[2]. For the waveguide, it is nature of the mode that exhibits hysteresis, being guided or radiated as described by the propagation constant β, and in this case the control parameter is the power flow down the guide, with the guide geometry taking the role of the order parameter.

In reality, the plane wave solution is not expected to be stable in the nonlinear medium[2]: however, hysteresis has been shown to exist for the bulk nonlinear interface[1], and so it might be anticipated that these results would offer a similar degree of qualitative agreement with experiment.

3. EXPERIMENTAL WAVEGUIDE DEVICE

Experiments were performed using carbon disulphide, CS_2, as the nonlinear medium. CS_2 has a high nonlinear coefficient and fairly well-established optical properties. The experimental device requires the waveguide surface in contact with the nonlinear medium to have a refractive index lower than the overlayer by an amount on the order 10^{-3} ideally, at low intensity excitation. Although some degree of tuning of the CS_2 index can be achieved by adjusting temperature or dilution, there remains the requirement that a material be identified capable of forming a waveguide of low propagation loss and refractive index matched to this index. Additionally, the waveguide must support only a single propagating mode (to avoid potential problems with intermode coupling) and be capable of withstanding the relatively high power densities demanded. All of these requirements can be met by utilising an optical waveguide formed by ion-exchange in glass.

Waveguides formed by exchanging monovalent cations in a glass substrate with cations of a different species in a molten salt are well-known[6]. In addition to their capability for low propagation loss, typically <0.5 dB.cm^{-1}, their surface refractive index can be varied continuously, with good control, by adjusting the concentration of the exchange cation in the molten salt, a property of particular significance for the present work. The ion-exchange technique produces a gradient distribution of refractive index in the depth of the glass substrate, with peak refractive index at the surface. The dispersion and modal properties of these waveguides are broadly similar to those of the uniform waveguide described in section 2, and the results presented there are valid in this case.

To facilitate measurement, a mode depth of a few μm is desirable, and to achieve this simultaneously with truly monomode behaviour requires a glass of refractive index ~1.56-1.59. We have used two glasses - Schott glass types LF5 and F5 with refractive indeices 1.5659 and 1.5864, respectively. Both glasses possess sufficient sodium to produce the required waveguides by ion-exchange of Na^+ with Ag^+ in molten $AgNO_3$. Waveguides were produced from molten $AgNO_3$ diluted with $NaNO_3$ to overcome problems associated with high silver ion concentrations at the surfaces of waveguides formed using the undiluted melt. Careful control of the $AgNO_3/NaNO_3$ ratio enables the resulting waveguide surface index to be adjusted to any specific value up to the maximum achievable using the undiluted $AgNO_3$ melt. Fig. 4 shows the measured surface refractive index variation with the melt concentration for both glasses, established using standard techniques[6].

The refractive index variation at 1.06 μm wavelength of CS_2 with temperature was measured using a hollow prism spectrometer: results are shown in Fig. 5, indicating a temperature coefficient of $-7 \cdot 10^{-4}$ K^{-1}. Temperature tuning of this index is used to make fine adjustments to the overlayer refractive index in the waveguiding experiments.

Fig.6 shows an experimental waveguide device. The CS_2 is confined in contact with the waveguide surface in a quartz cell bonded to the waveguide using a CS_2-resistant epoxy resin. The liquid CS_2 is injected into the cell through a vapour-tight PTFE seal mounted in the cell top wall, to prevent evaporation loss. Light is injected into the ion-exchanged glass waveguide by prism-coupling(7) before encountering the nonlinear overlayer. Care is taken in forming the epoxy contact between the cell and the waveguide to introduce a gentle taper into the epoxy profile to minimise loss of guided optical power at the transition. The entire assembly is mounted in a temperature-controlled enclosure, designed to permit light input and output coupling and maintain the CS_2 temperature within better than $\pm 0.1^{\circ}C$ of a preset value. The partially assembled waveguide cell enclosure is shown in Fig. 7.

4. EXPERIMENTAL RESULTS

Experiments were performed by prism coupling the output from a Q-switched Nd-YAG laser generating 25 mJ pulses of 15 ns duration at a wavelength of 1.06 μm. This beam was focussed to produce a ~1 mm wide beam in the waveguide beneath the cell containing the CS_2 and attenuated by a range of optical density filters. The waveguide supported only the fundamental TE_0 mode in all regions under low power excitation.

The experimental arrangement is shown in Fig. 8. A beam splitter and optical delay is used to split off a reference pulse from the main beam which is detected and displayed time-displaced from the signal pulse on an oscilloscope. We choose not to monitor the guided wave that emanates from the end of the device as there is considerable light scattered foward in this general direction including that radiated from the guided mode through the nonlinear interaction: it is generally difficult to spatially isolate only the guided mode. Instead we derive the signal pulse from the radiation scattered from the waveguide in the interface region. Thus, light scattered and coupled from the waveguide leaves the device at angles that are characteristic of the modes of propagation of the structure. Light scattered from the guided mode is monitored and the detected signal displayed on an oscilloscope. An oscilloscope trace obtained in this manner is given in Fig. 9, showing the signal and reference pulses.

Although the pulses available from the laser are complicated by possessing some structure due to the presence of several longitudinal modes, it is clear from Fig. 9 that the signal pulse is somewhat narrower and more asymmetric than the reference pulse. The pulses are replotted superimposed in Fig. 10, demonstrating these effects more clearly. The effect can be removed completely by temperature tuning the CS_2 refractive index to a lower value such that the input power density is insufficient to achieve $n_3 > n_1$.

The data derived from this experiment can be replotted to show the signal pulse intensity as a function of the input pulse intensity. This is shown in Fig. 11 which indicates the presence of hysteresis in the response. We have achieved similar response for a range of peak powers coupled into the structure and find the effect to be present above a certain critical intensity as expected from the theory given in section 2. It is somewhat difficult to accurately measure the threshold intensity, which is predicted to be $\sim 1\text{-}2 \times 10^8$ Wcm^{-2} by the plane-wave theory for the waveguides used. Allowing for the coupling loss and modal mismatch loss at the linear/nonlinear waveguide transition, we estimate the threshold intensity to be not greater than $\sim 10^9$ $W\,cm^{-2}$.

We have found that this effect is repeatable although the input pulse structure does complicate interpretation of the measurements: we are now modifying the laser resonator to restrict oscillation to a single longitudinal mode. Further effort is now required to quantify the process and make time-resolved measurements.

5. CONCLUSION

Intensity-dependent hysteresis has been observed in the power guided by an optical waveguide comprising an ion-exchanged guiding layer in glass surmounted by CS_2. The CS_2 exhibits a large positive optical Kerr coefficient providing the strongly intensity-dependent refractive index that is responsible for the observed behaviour. A straightforward plane-wave theory gives qualitative agreement with the observed behaviour. This effect should form the basis for a range of fast optical processing devices.

6. REFERENCES

(1) P W Smith, W J Tomlinson, P J Maloney and J P Hermann, "Experimental studies of a nonlinear interface", IEEE J Quantum Electron, QE-17 (3), March 1981, pp 340-348.

(2) A E Kaplan, "Theory of hysteresis reflection and refraction of light by a boundary of a nonlinear medium", Sov Phys JETP, 45, May 1977, pp 896-905.

(3) P W Smith and W J Tomlinson, "Nonlinear optical interfaces: switching behaviour", IEEE J Quantum Electron, QE-20 (1), Jan 1984, pp 30-36.

(4) D J Robbins, "TE modes in a slab waveguide bounded by nonlinear media",Optics Commun, 47 (5), Oct 1983, pp 309-312.

(5) G I Stegeman, "Guided wave approaches to optical bistability", IEEE Trans Microwave Theory Techn, MTT-30 (10), Oct 1982, pp 1598-1607.

(6) G Stewart, C A Millar, P J R Laybour, C D W Wilkinson and R M De la Rue, "Planar optical waveguides formed by silver-ion migration in glass", IEEE J Quantum Electron, QE-13 (4), April 1977, pp 192-200.

ACKNOWLEDGEMENT

The authors wish to acknowledge useful discussions with A Miller, RSRE Malvern, UK. This work has been carried out with the support of the Procurement Executive, Ministry of Defence.

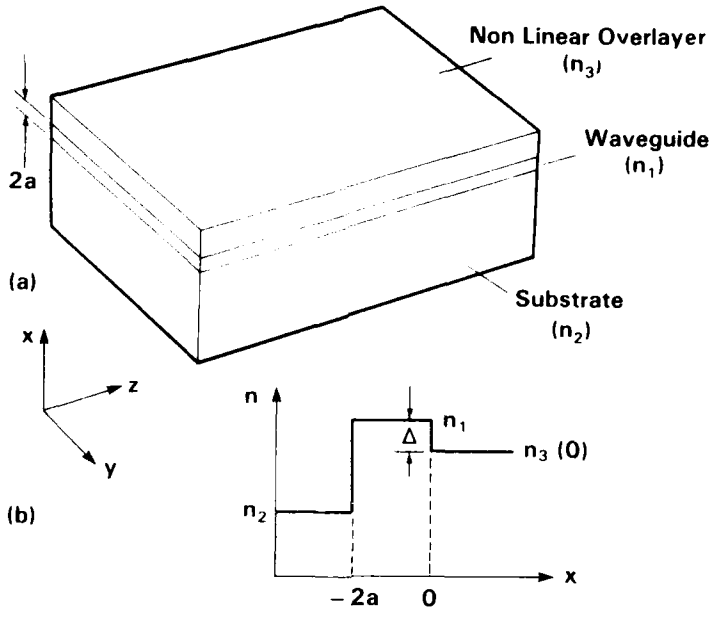


FIGURE 1 (a) PLANAR OPTICAL WAVEGUIDE WITH NONLINEAR OVERLAYER
(b) RELATIVE REFRACTIVE INDICES OF THE MEDIA

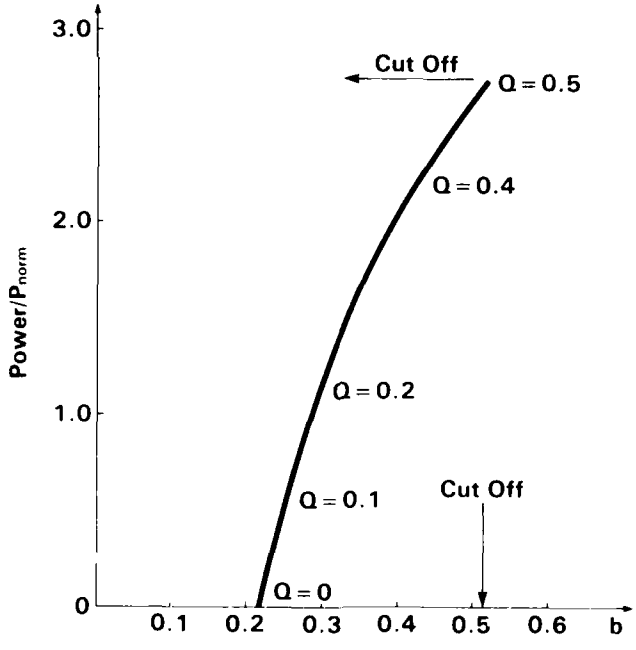


FIGURE 2 NORMALISED POWER FLOW P/P_{norm}, AS A FUNCTION OF NORMALISED PROPAGATION CONSTANT, b, FOR A WAVEGUIDE WITH d = 10, v = 1

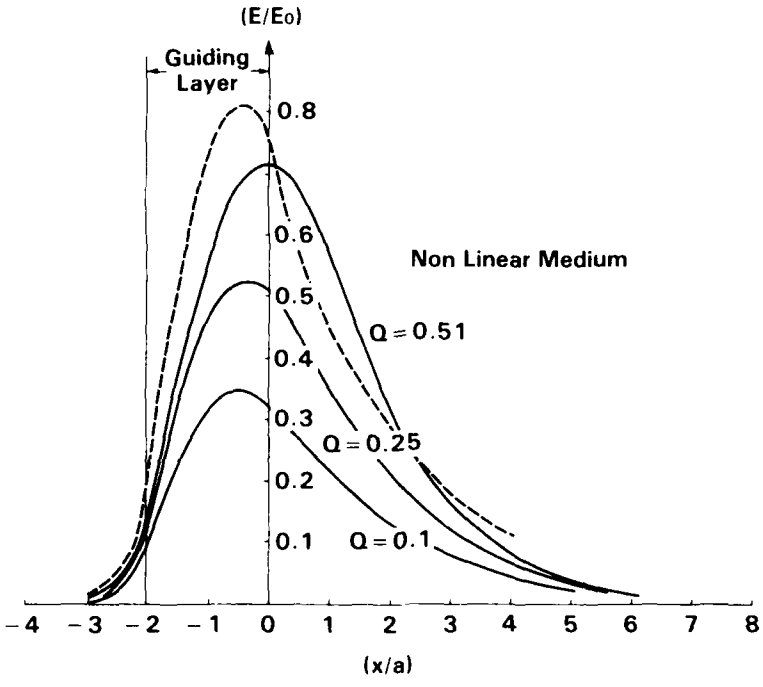


FIGURE 3 MODE PROFILES FOR WAVEGUIDE WITH d = 10, v = 1. THE DASHED CURVE SHOWS A CORRESPONDING MODE PROFILE FOR A LINEAR GUIDE NORMALISED TO THE SAME INTERFACES

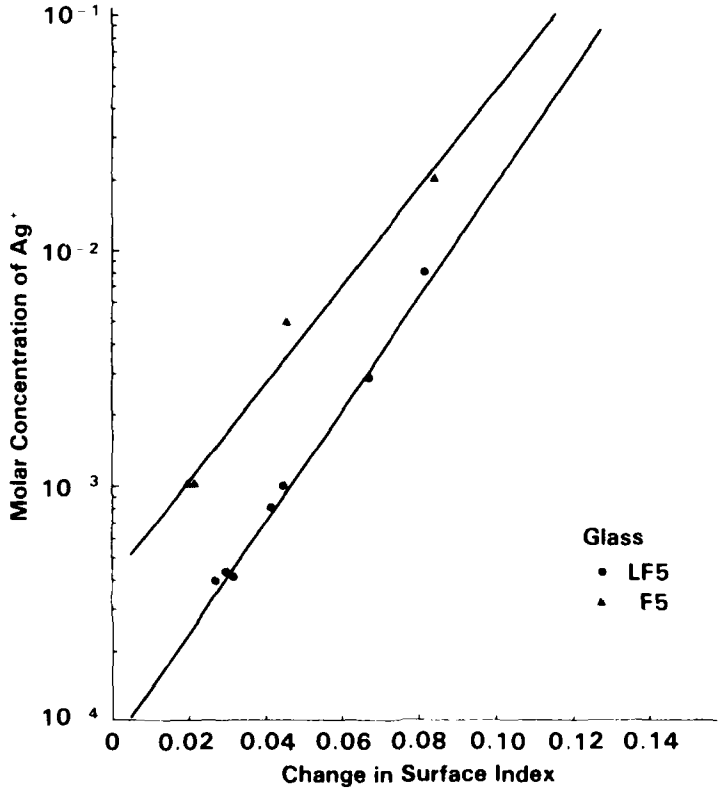


FIGURE 4 REFRACTIVE INDEX CHANGES AS A FUNCTION OF MOLAR CONCENTRATION FOR ION-EXCHANGED GLASS WAVEGUIDES

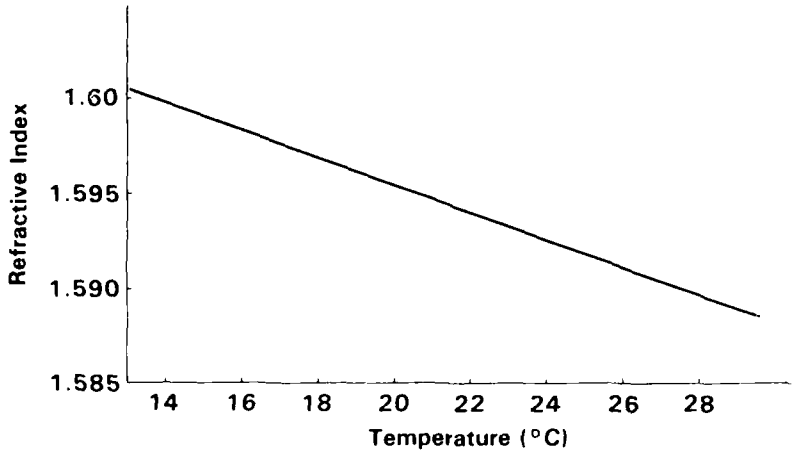


FIGURE 5 REFRACTIVE INDEX OF CARBON DISULPHIDE AT 1.06um AS A FUNCTION OF TEMPERATURE

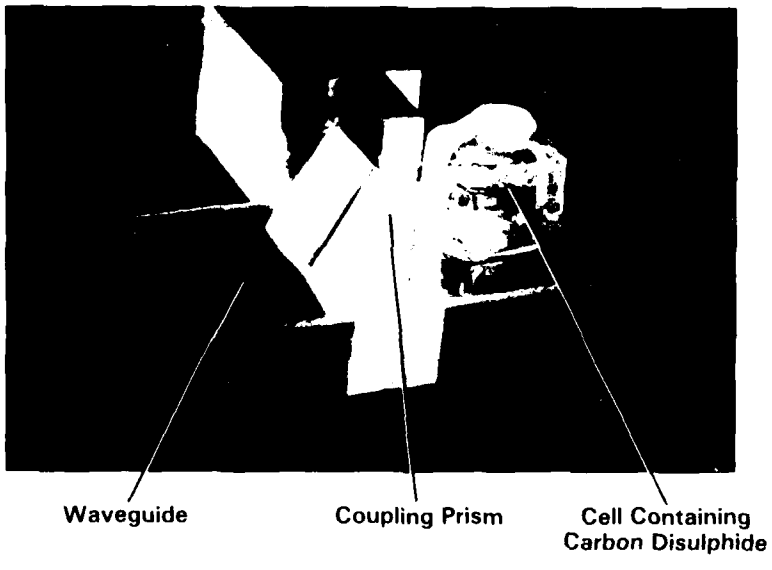


FIGURE 6 EXPERIMENTAL WAVEGUIDE-NONLINEAR INTERFACE DEVICE

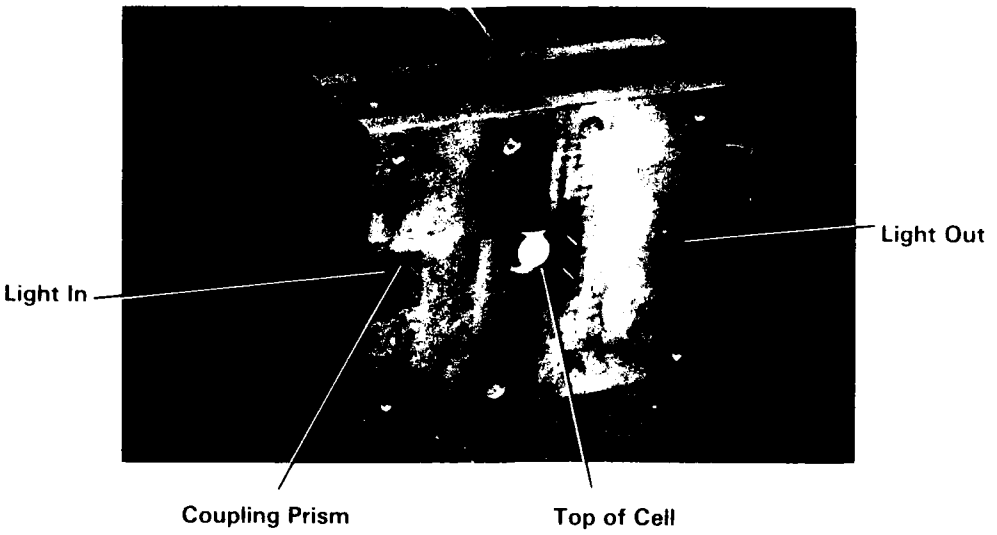


FIGURE 7 PARTIALLY ASSEMBLED WAVEGUIDE-CELL THERMAL ENCLOSURE

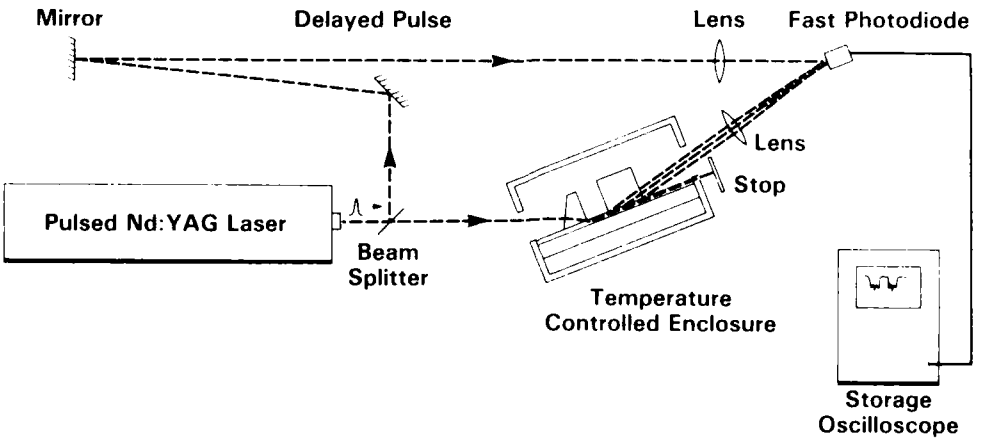
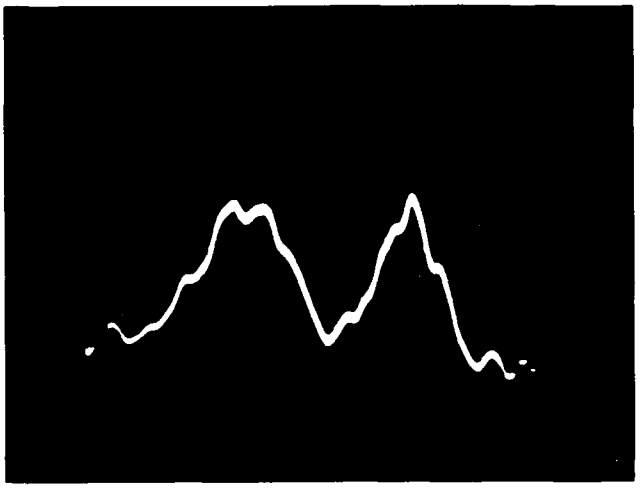


FIGURE 8 EXPERIMENTAL CONFIGURATION FOR DEVICE ASSESSMENT



Reference Pulse Signal Pulse

FIGURE 9 OSCILLOSCOPE TRACE SHOWING INPUT AND OUTPUT PULSES

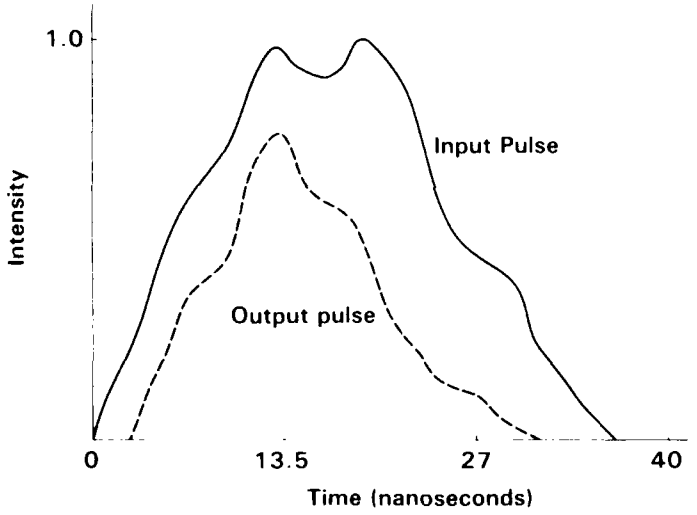


FIGURE 10 MEASURED INPUT AND OUTPUT PULSE FORMS

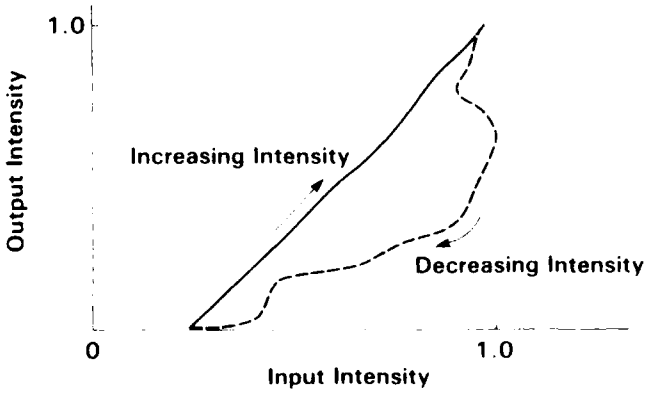


FIGURE 11 EXPERIMENTAL PULSE RESPONSE REPLOTTED TO SHOW HYSTERESIS

DISCUSSION

R.G.Harrison, UK

You're talking about CS_2 and the fact that the nonlinearity is not sufficiently high, but are you aware that there are some reports of resonant enhancement using 10 micron radiation for CS_2 -- back a decade or two I seem to recollect a paper on self-focusing using multikilowatts rather than megawatts

Author's Reply

Yes, I think I have to come in straight away and say that a lot of our interest has been to try to hold the wavelength near to our other interest. My main intetests here have always been to see these things as having fibres associated with them somewhere. We tend to go in saying this must look, or try to look, as near to the existing technology as we have. We do a lot of work therefore, between 1.3 and 1.55 microns. 1 micron is near enough -- 10 is too far away at the moment

J.R.Forrest, UK

When the wave enters this region how much of the energy is reflected back? Obviously you've been talking about the scattering out of the layer but there must presumably be some reflection back associated with the discontinuity. Could you give us some idea of the magnitudes on that?

Author's Reply

Yes. Typically, depending once again about how clever we've been about making the little taper, there can be as much 50% of the energy scattered back. There is also leakage on the other side, so that light is in fact leaked out by the presence of the taper also. One can do the usual thing and go around and add up all the intensity around the device. I think that typically between 10 and 25% of the energy actually reaches the nonlinear area of the guide and is useful at that point -- perhaps even between 5 and 25%.

H.M.Gibbs, US

Would you summarize the differences between your theory and those of Stegeman and of Sarid?

Author's Reply

I think that the main difference is that ours is incomplete. My feeling is that I'm not sufficiently familiar yet with what George Stegeman has done. I think that it does look as though ours is converging towards his and the answer is indeed the same. I must confess that at first I was suspicious of that theory, but it does look as though our solution is going along that way. We've also been doing some work with Alan Boardman who has been working or talking to Stegeman about this; it is coming together and it is his theory that looks right

$$E_i = E_{ex} + \sum_{j \neq i} A_{ij} P_j(t) \qquad (17)$$

where E_{ex} is the amplitude of the externally-applied field, P_j is the dipole amplitude of atom j and

$$A_{ij} = \frac{e^{ikr_{ij}}}{kr_{ij}} \left[1 + \frac{i}{kr_{ij}} - \frac{1}{(kr_{ij})^2} \right] . \qquad (18)$$

In this equation, k is the magnitude of the wave vector, and r_{ij} is the distance from atom i to atom j. For simplicity and to limit the number of parameters $r = r_{ij}$, and $A = A_{ij}$ are taken to be independent of i and j.

The role played by A has two essential features. The Re A causes a frequency shift in the resonance condition causing OB and varies strongly with separation r. We find that OB occurs for the condition

$$B = (N-1) \text{ Re } A = -6 . \qquad (19)$$

The Im A gives rise to superradiant and subradiant decays and modifies the bistable behavior, but does not affect the bistable condition which is (19). It varies slowly with r for $kr \gg 1$ and is taken to be constant in this calculation.

The calculation is performed numerically and the results are treated as experimental data. At every step in the calculation the random number generator determines whether an atom will be restarted in the ground state, independently of what state it is in at that time. On the average, an atom is restarted in the ground state once in one unit of time, that unit being the decay time. Retardation is included exactly in all phase factors, but is neglected in the slowly-varying amplitudes of the dipoles. No prior knowledge is available as to what variables, if any, are bimodal. The meaning of upper and lower state is arbitrary.

Let $n_i(t)$ denote the inversion of atom i. Let

$$\bar{n}(t) = \frac{1}{N} \sum_{i=1}^{N} \frac{1}{T_f} \int_{-\infty}^{t} dt' \, e^{(t-t')/T_f} \, n_i(t') \qquad (20)$$

be the inversion averaged over all the atoms, observed with a detector with time constant T_f, i.e., the detector has bandwidth $1/T_f$. Let $\bar{n}_c$ be a constant to be determined later, such that if

$$\bar{n} > \bar{n}_c \qquad \text{system is in "upper state"} , \qquad (21a)$$

$$\bar{n} < \bar{n}_c \qquad \text{system is in "lower state"} . \qquad (21b)$$

An "up" ("down") time, t_u (t_d) is defined as an interval from the time the system enters the up (down) state to the time it leaves it. We will denote T_u (T_d) as the average up (down) time.

If the system is responding according to a single statistically stationary process, then the distributions $P(t_u)$ and $P(t_d)$ of t_u and t_d are exponential,

$$P(t_u) = \frac{1}{T_u} e^{-t_u/T_u} , \quad P(t_d) = \frac{1}{T_d} e^{-t_d/T_d} \qquad (22)$$

for which the variances of t_u and t_d must equal the mean. This can be used as a consistent check on whether T_f and $\bar{n}_c$ are properly chosen in the calculation. When bimodality is observed, we find that

$$\bar{n}_c = (n_1 T_2 + n_2 T_1)/(T_1 + T_2) \qquad (23)$$

So, the effect of the local field correction is to cause an inversion-dependent frequency shift in the equations of motion, Eqs. (10) and (11). This set of equations is quite similar in form to the equations of motion for the quantized driven, damped Duffing oscillator[14]. The only difference is the appearance of the factor of n in the last term on the right-hand side of Eq. (11). Since OB is known to exist for the classical, driven, damped Duffing oscillator[4] in the absence of a cavity (i.e., feedback of the electromagnetic field), it is then also expected in this case under very much the same circumstances. This will depend crucially upon the initial detuning δ, as well as the absorption, i.e., B.

The atomic equations (10) and (11) are solved in steady state in terms of R_{ab} by eliminating n, and the result is used in Maxwell's equation (2) in the plane wave, slowly-varying amplitude (in time) limits and using the relation (5). The boundary conditions are that, at $z = 0$,

$$\mathcal{E} = \mathcal{E}_R + \mathcal{E}_I \quad , \qquad \frac{d\mathcal{E}}{dz} = ik(\mathcal{E}_I - \mathcal{E}_R) \quad , \tag{13a,b}$$

and at $z = L$

$$\frac{d\mathcal{E}}{dz} = ik\mathcal{E} \quad . \tag{14}$$

Here, $\mathcal{E}_R$ and $\mathcal{E}_I$ are the reflected and incident fields, respectively, and k is the wave vector for the incident field. The details of the calculations are presented in reference 10, and will not be repeated here.

The main results are shown in Figures 1 and 2, which show OB in steady state for the transmittance, T, as a function of the initial detuning δ for various values of the incident field E_I and the absorbance A as a function of the incident field E_I, respectively. The transmittance T and reflectance R are defined in the usual way by

$$T = |\mathcal{E}_T|^2 / |\mathcal{E}_I|^2 \quad , \qquad R = \frac{|\mathcal{E}_R|^2}{|\mathcal{E}_I|^2} \quad , \tag{15}$$

and the absorbance A,

$$A = 1 - R - T \tag{16}$$

is the fraction of optical energy dissipated in the medium. From Figure 2, it is seen that the intrinsic OB is absorption-dominated, i.e., the reflection is small and absorbance is the complement of the transmittance. This, together with the inversion-dependent frequency shift nonlinearity, Eq. (11) and initial detuning dependence are all common characteristics of intrinsic mirrorless OB due to absorption[5-8,10].

III. SEMICLASSICAL MICROSCOPIC MODEL

To verify the validity of using the LFC as a cause of cavityless OB analyzed in the previous section, and at the same time determine the effect of fluctuations on the conditions for OB, using the LFC, we consider in this section a microscopic model where the local field drives each atom and therefore there is no LFC to be made. If the results of the previous section are physical, then the essential features must appear in the results from the microscopic model. That is, if the LFC is, indeed, a correction necessitated by passage from the microscopic to the macroscopic representation[9], then, certainly, the essential features should be intrinsic in the microscopic counterpart. In addition, we analyze the effect of fluctuations.

The microscopic model presented here is a semiclassical heuristic one with fluctuations. In this treatment, we draw from the analogy between bistability and first-order phase transitions in that hysteresis phenomena are known to be subtle features of stochastic models. It is useful therefore to compare essential features of deterministic models as opposed to stochastic models used to describe first-order phase transition phenomena.

In deterministic models, one examines the mean of some variable that has nonunique values, whereas stochastically, one predicts a distribution for the variable. Deterministically, then, one predicts several stable states and hysteresis, where on the other hand, the distribution function for stochastic models has a unique first moment, and therefore no hysteresis for the corresponding mean variable. Hysteresis in stochastic models, on the other hand, is contained in a bimodal distribution function. Thus, a higher-order moment analysis is needed to expose the bimodal condition. Thus, the contrast between deterministic models representing a first-order phase transition and stochastic models is that deterministic models give conditions for bistability, whereas stochastic models give rise to bimodality.

Our model consists of a limited number of two-level atoms which obey decayless Bloch equations. The field amplitude E_i at atom i is given by

$$\nabla^2 \underline{E} + \frac{1}{c^2}\frac{\partial^2 \underline{E}}{\partial t^2} = \frac{4\pi}{c^2}\frac{\partial \underline{P}}{\partial t} \quad . \tag{2}$$

Here, $\underline{P}$ is the macroscopic polarization

$$P = \sum_j P_j \quad . \tag{3}$$

The major point is that when the field in (2) is used to drive atom m, it contains the self-field of that atom. As shown by Van Kronendonk and Sipe[9] (VKS), the self-field contribution is removed by taking the atom to be driven by the "local field", E_L,

$$E_L = E + \left[\frac{4\pi}{3} - s\right] P \quad , \tag{4}$$

which was derived for homogeneous-broadening and cubic or spherical local symmetry in the arrangement of the atoms if $s = 0$ and s is otherwise a structure factor. The structure factor stems from the correlation of atoms that occurs in crystals[12].

For the case we treat here, the slowly-varying envelope approximation[13] (SVEA) is made only in the time variable t, not in the spacial variable z, since as we shall show, OB occurs in a small volume in this case only for a dense medium, so the field varies strongly spacially and the SVEA spacially is totally invalid. In terms of slowly-varying amplitudes (temporally), the polarization P and the off-diagonal density-matrix element R_{ab} for a two-level atom, are related by

$$P = i u N R_{ab} \quad . \tag{5}$$

Here, u is the matrix element of the transition dipole moment and N is the density of atoms. The slowly-varying amplitude of the local field, $\mathcal{E}_L$, is

$$\mathcal{E}_L = \mathcal{E} + i\pi u N R_{ab} \quad , \tag{6}$$

where $\pi = \frac{4\pi}{3} + s$. The density matrix equations for two-level atoms with upper (lower) states labeled a(b) are

$$\frac{dR_{ab}}{dt} = -(i\Delta + \gamma_{ab})R_{ab} + \frac{u}{2\hbar}\mathcal{E}_L n \quad , \tag{7}$$

$$\frac{dn}{dt} = -\gamma(1+n) + \frac{u}{\hbar}(\mathcal{E}_L^* R_{ab} + c.c.) \quad , \tag{8}$$

where the inversion n is

$$n = \rho_{aa} - \rho_{bb} \quad . \tag{9}$$

If the local field $\mathcal{E}_L$ is eliminated from (7) and (8) using (6), the result is

$$\frac{dn}{dt} = -\gamma(1+n) + \frac{u}{\hbar}(\mathcal{E}^* R_{ab} + c.c.) \quad , \tag{10}$$

$$\frac{dR_{ab}}{dt} = -[(i\Delta - B\gamma_{ab} n) + \gamma_{ab}]R_{ab} + \frac{u}{2\hbar}\mathcal{E} n \quad , \tag{11}$$

where

$$B = \frac{\lambda^2 \alpha}{16\pi^2} \tag{12}$$

and α is the absorption coefficient and λ is the wavelength of light in vacuum.

CAVITYLESS OPTICAL BISTABILITY IN SYSTEMS OF TWO-LEVEL ATOMS

C. M. Bowden
Research Directorate, US Army Missile Laboratory
US Army Missile Command
Redstone Arsenal, Alabama 35898-5248, USA

and

F. A. Hopf
Optical Sciences Center
University of Arizona
Tucson, Arizona 85721, USA

SUMMARY

The local field correction (LFC) is used to eliminate the local field in the macroscopic semiclassical density matrix for a collection of homogeneously-broadened, two-level atoms interacting with the radiation field, which is described by Maxwell's equation. The system of equations is solved in steady state with the boundary conditions for a plane wave normally incident on the dense absorbing medium of propagation length much less than a resonance wavelength. Optical bistability (OB) is predicted, based upon adiabatic or steady-state conditions. As a test of the validity of the predictions, results are reported using a semiclassical heuristic microscopic model with fluctuations for a collection of a finite number of electric field-driven, two-level atoms. In this case no LFC is needed and none is made; however, OB is observed consistent with the predictions from the steady-state macroscopic model.

I. INTRODUCTION

Mirrorless optical bistability (OB) was predicted by Bowden and Sung[1] and by Bowden[2,3], using a fully-quantized model, for a collection of two-level atoms in a small volume, driven by an externally-applied laser field. A first-order phase transition, i.e., switching, was predicted due to interatomic correlation via the electromagnetic field. Another suggestion for cavityless OB was discussed by Flytzanis and Tang[4], who used a classical model based on the driven, damped Duffing oscillator. Recently, Hajto and Janossy[5] observed OB in amorphous $GeSe_2$ due to increasing temperature-induced narrowing of the bandgap. More recently, Miller[6] and co-workers observed mirrorless OB in a GaAs/GaAlAs quantum-well semiconductor caused by increasing absorption. Similar results were observed by Dagenais and Sharfin[7] in CdS. Conditions for mirrorless OB based upon nonlinear absorption were proposed by Henneberger and Rossmann[8].

To further systematically investigate causes and conditions for cavityless OB we have treated the case for the effect of the local field correction (LFC)[9] in a macroscopic semiclassical model of two-level atoms driven by a local field[10]. Although of completely different mechanism than any of the cases cited above, the OB which is predicted is characterized by large absorbance, i.e., the state of large absorption corresponds to low transmission, just the opposite of the conditions for high-Q cavity OB[11]. In the development of our model[10], we draw heavily upon the work of Van Kronendonk and Sipe[9] and their interpretation of the origin of the LFC as the necessity for removal of an atom's self-field from the macroscopic field in the interaction, in passage from the microscopic to the macroscopic semiclassical models. We find that resonatorless OB is possible from the macroscopic semiclassical approach[10], due to the LFC[9] for particular conditions of detuning of the incident field from the atomic transition for a suitably dense medium of two-level atoms.

The purpose of this paper is to present results of calculations for mirrorless OB derived from a microscopic statistical model in which the LFC is not needed[9] and therefore is not made. This is done to test the results of the macroscopic model and, as we show, the results confirm the existence of resonatorless OB in a collection of two-level atoms driven by an externally-applied field. Our results suggest that experiments be conducted on suitable materials to study the phenomenon which is both of profound fundamental significance as well as potentially useful in the applied sense.

The next section will be used to present our macroscopic model[10] and to briefly discuss the main results for mirrorless OB. Then in Section III we present the model and main results of our microscopic, statistical treatment. The final section is used for discussion and summary.

II. MACROSCOPIC MODEL AND LOCAL FIELD CORRECTION

In the microscopic theory of electromagnetism, the force on an atom, labeled by the integer m, is due to the field, denoted by E_m,

$$E_m = E_1 + \sum_{\substack{j=1 \\ j\neq m}}^{N} G_{mj} P_j \qquad (1)$$

where E_1 is the incident field, the P_j are the polarizations of atoms j, $j = 1,2,\ldots,N$, $j\neq m$. The Green's functions G_{mj} are the Lenard-Weichart potentials and take into account the geometry. The absence of m in the sum in Eq. (1) stems from the removal of the self-field of atom m and appears in the theory as natural relaxation and spontaneous emission.

In the usual passage to macroscopic electromagnetism the self-field is reintroduced in the form of Maxwell's equation, which for isotropic media, is

H.M.Gibbs, US

Spectroscopists dislike Fabry-Perot fringes; they try very hard to avoid them and often AR-coat samples. So if you look at spectroscopy experiments, you should not see Fabry-Perot fringes. There are two experiments on copper chloride using etalons: Levy's group, Honerlage, etc. in Strasbourg and ours; both of us go to great pains to make sure that we have Fabry-Pero fects.

Author's Reply

But again in these experiments which I was referring to there's never any word about antireflecting coatings. So I'm up in the air — I don't know if they did it and didn't say anything — what would be the explanation?

C.Klingshirn, Ge

Just a comment: the quality of the surfaces of the copper chloride platelets is too bad to see Fabry-Perot modes, and therefore you may have to do artificial things to get at least a little bit of Fabry-Perot modes. Other materials, like cadmium sulphide, grow with perfect Fabry-Perot modes without any additional coating or something else.

DISCUSSION

C.Klingshirn, Ge

I have a question concerning the group velocity. You eliminate from your calculation the spatial dispersion by assuming for the exciton and biexciton an effective mass of infinity as compared to the actual values of about 2.5 m_0 and 5 m_0, respectively; this disregards the k dependence of the exciton resonance of the biexciton and with this also the k dependence of the induced resonance around half the biexciton resonance. Now if you include this, this greatly alters the dispersion and especially the group velocity. Have you any ideas how this would influence the switching times?

Author's Reply

In the neighbourhood of the biexciton resonance of course the group velocity would be much different from that calculated from that expression — it'll be even lower, and so I expect even longer switching times at the biexciton resonance. Now the optical bistability has only been observed near the resonance, and at the moment I have no explanation of why that's the case. We're trying to think of ideas why it's only been observed very close to the biexciton resonance; we think it should be there far away from resonance. The original proposal of Hanamura was to get away from this biexciton resonance and the exciton resonance, so that you only get virtual formation of biexcitons, and therefore you'll get a very fast switch; that's what we were here attempting to study.

S.D.Smith, UK

Is there any reason to suppose that the biexciton nonlinearity will be actually more efficient than that which you might get in a semiconductor by saturating an exciton or a band edge or inducing a plasma and then trimming the carrier lifetimes to fit to such a reaction time? Might the system actually give you a better ratio of nonlinearity to absorption?

Author's Reply

No, I don't believe it will be better. In fact we've been using this as a sort of proving ground for approximations to try to develop ways of handling the propagation in the problem to include these other effects. So it's not really the problem we want to solve; we want to go on to things like gallium arsenide and the multiple-quantum-well structures and use the same sort of techniques there that we've developed here. Those are much more complicated models to solve in that case, so we want to first understand what's happening in copper chloride and move on to gallium arsenide, or maybe do both at the same time to attack these problems from the point of view of the propagation and switching times. We're interested in the transient dynamics.

H.M.Gibbs, US

A comment on what Professor Smith just said. In CuCl the binding energy is very high for the exciton so the saturation intensity is extremely large, so I think that you would expect the nonlinearity to be larger on the biexciton in that particular system. But that doesn't make it more attractive that excitons that have much lower saturation intensities in other systems. A question for Dr Haus: I'm not sure what you mean when you say that the sensitivity to length should be considered in experiments; I think that that sensitivity is well know by experimentalists — if you mean some fraction of an instrument width of a Fabry-Perot.

Author's Reply

That's precisely what I mean. The quality of the surfaces has never been discussed in the literature. I'd like to have that problem understood. Even with just the transmission of light through a thin sample or a thick sample, there is never any mention of the surface quality. In fact I spent several telephone conversations with Lloyd Chase in Indiana in the past couple of months. According to him they just grow these crystals, and they look for a spot at which they get good transmission, but they can't really say that the surfaces are high quality. That's what I get from the experimentalists.

H.M.Gibbs, US

One has some measure of surface quality because one can measure the experimental finesse; one has some indication of the difference between the reflecting finesse and the actual observed finesse, and in most cases that's due to nonflatness. You can tell from the wedge whether that's a general gradual nonflatness or whether it has to do with local surface irregularities.

Author's Reply

Yes, but in the transmission experiments we expect to see these etalon effects and they're not there in the experiments.

H.M.Gibbs, US

Whose experiments?

Author's Reply

Well if you look at the Japanese experiments by Shionoya, I believe? And the experiments by Chase and another Japanese group. The only experiments where I've seen the etalon effects are in these new optical phase conjugation experiments of Lloyd where he gets oscillations in the intensity as he sweeps across the frequency.

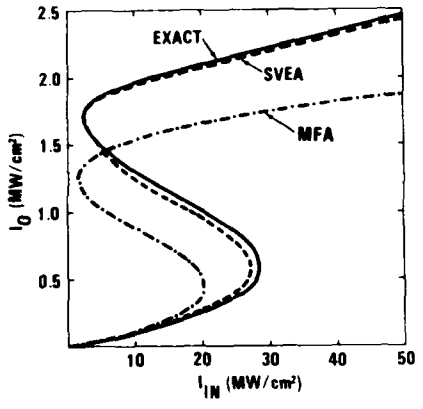


Figure 1. Input versus output intensities for the laser frequency ω = 3177 mev and sample length L = 9.97753 μm. Other relevant parameters are: R = .9, γ_m = 0.3 mev and γ_x = 0. The calculation for the second-order Maxwell equation (solid line) is compared to the mean-field approximation (-.-.-.) and slowly-varying envelope approximation (-----).

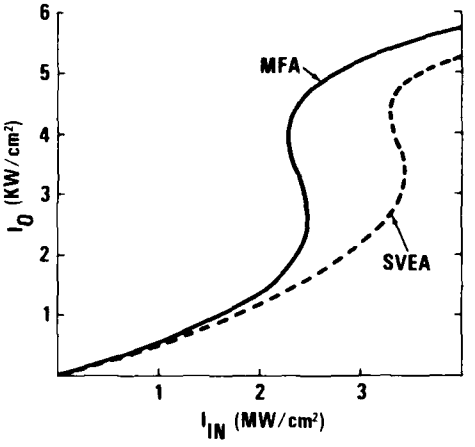


Figure 2. Input versus output intensities for the laser frequency ω = 3186.2 mev and L = 30 μm. The solid line corresponds to the MFA and the dashed line is the SVEA. The remaining parameters are: R = .9, γ_m = 0.04 mev and γ_x = 0.

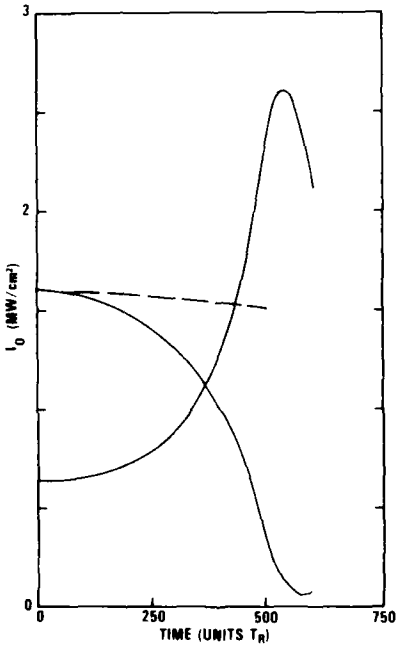


Figure 3. Output intensities versus round trip time T_R = .165 psec for switch-up and switch-down situations. The same parameters are used as in Figure 1, however, here γ_x = .03 mev.

TABLE

MATERIAL CONSTANTS

$\varepsilon_\infty = 5$

$4\pi g_1^2 = 27.5$ mev

$M^2 = 1.57 * 10^{-16}$ mev^2 - cm^2

$g_2^2 = \gamma_R(0) * M^2/\ldots$

$\omega_x = 3202.7$ mev

$\omega_m = 6372.5$ mev

REFERENCES

1. H. Haug, R. März, and S. Schmitt-Rink, Phys. Lett. 77A, 289 (1980); R. März, S. Schmitt-Rink and H. Haug, Z. Physik B40, 9 (1980); S. W. Koch and H. Haug, Phys. Rev. Lett. 46, 450 (1981); E. Hanamura, Solid State Commun. 38, 939 (1981).

2. C. C. Sung and C. M. Bowden, "Optical Bistability in Semiconductors", in Optical Bistability II, eds., C. M. Bowden, H. M. Gibbs and S. L. McCall, (Plenum, New York, 1984), p. 241.

3. C. C. Sung and C. M. Bowden, Phys. Rev. A29, 1957 (1984).

4. D. Sarid, N. Peyghambarian and H. M. Gibbs, Phys. Rev. B28, 1184 (1983).

5. C. C. Sung and C. M. Bowden, J. Opt. Soc. Am. B1, 395 (1984).

6. C. C. Sung, C. M. Bowden, J. W. Haus and W. K. Chiu, Phys. Rev. A., 1984, to be published.

7. N. Peyghambarian, D. Sarid and H. M. Gibbs, in Reference 2, p. 259; N. Peyghambarian, H. M. Gibbs, M. C. Rushford and D. A. Weinberger, Phys. Rev. Lett. 51, 1692 (1983).

8. B. Hönerlage, J. Y. Bigot and R. Levy, in Reference 2, p. 253.

9. U. Geigemüller, U. M. Titulaer and B. U. Felderhof, Physica 119A, 41 (1983); F. A. Haake, Z. Phys. B48, 31 (1982).

10. A. Icsevgi and W. E. Lamb, Phys. Rev. 185, 517 (1969).

11. J. A. Fleck, Phys. Rev. B1, 84 (1970).

12. F. P. Mattar, G. Moretti and R. E. Granceour, Computer Phys. Comm. 23, 1 (1981).

13. S. L. McCall, Phys. Rev. B9, 1515 (1979).

14. See Hanamura's paper in Reference 1.

15. J. W. Haus, C. M. Bowden and C. C. Sung, to be published, (1984).

We have chosen ω = 3177 mev (i.e., Δ = 15.5 mev) so that the absorption is small. Also, r_m = .9 has been chosen to reduce the contribution from absorption and enhance the dispersive contribution. For these values, the EX, SVEA and MFA give qualitatively identical results. Quantitatively, the SVEA is well within 5 percent of the EX results, whereas the MFA curve varies as much as 30 percent from the other OB curves. The remarkable agreement between the SVEA and EX results is due to the choice of the wave vector $\hat{k}$, Eq. (17), and the fact that the nonlinearity which causes OB in CuCl is found to be very nearly of the Kerr type (more will be discussed concerning this behavior later). These conditions render the envelopes in (16) and (18) truly slowly-varying in the medium.

In all our numerical calculations we have used a predictor-corrector method. For the results in Figure 1, 40000 points were used to generate the EX results. Using only 20000 points resulted in 10 percent change in the curves. On the other hand, 1000 points were more than sufficient in obtaining the SVEA results. Our dynamical calculations reported below were performed with only 100 points and the steady-state curves for this grid differs from the above by about 3 percent.

A value for the incident field frequency ω is taken just below (0.2 mev) the biexciton two-photon resonance, and the results are presented in Figure 2 for L = 30 μm. To achieve a condition for OB, it was necessary to use a much smaller value for γ_m than for the previous cases. If OB in CuCl is strongly sensitive to tuning near the two-photon resonance, as has been suggested[1], and might be anticipated from the field dependence of the denominator in Eq. (9), we would expect to observe a large hysteresis in Figure 2, as compared to the previous cases. As observed, there is only weak OB indicated. Again, the SVEA and MFA results are in qualitative agreement, but quantitatively there is, as for the other case treated, a discrepancy of almost 30 percent between them.

The absence of sensitivity of OB in tuning near the two-photon biexciton resonance strongly suggests a Kerr type behavior of the nonlinearity, contrary to what has been previously suggested[1]. For further discussion on this matter, the reader is referred to reference 6.

The set of Eq. (18) and (22a-c) were integrated numerically without adiabatic elimination with γ_x = .03 mev, otherwise conditions were identical to those given in Figure 1. The dynamical behavior of the system near the turning points for the SVEA curve of Figure 1 is depicted in Figure 3 as the input field is slowly increased[15]. As indicated, the switch-down time is approximately 80 psec and is approximately equal to the switch-up time. This is in direct contrast to the prediction of reference 14 which predicts a switching time on the order of picoseconds. This discrepancy stems from the use of adiabatic elimination of the exciton, biexciton and exciton-biexciton variables. The conditions for adiabatic elimination[9] of atomic variables were not satisfied for OB in CuCl. The switching times correlate very closely with the polariton lifetime in the material. Since the material is highly dispersive, the group velocity of the polaritons is about 12 times smaller than the phase velocity at ω = 3177 mev. The group velocity determines the relaxation time of the cavity and is the dominant slow mode of the system. This effect was not accounted for in previous work[14]. Furthermore, we can infer that the switching times will become dramatically longer closer to the exciton resonance because the group velocity is inversely proportional to ξ^2. Using this argument at ω = 3186.2 mev and L = 10 μm, the switching times will be about 4 times longer and at ω = 3196 mev, the switching times will be about 25 times longer.

IV. CONCLUSIONS

Predictive results for model calculations in OB for excitonic-biexcitonic CuCl have been presented in this paper and comparative analysis was made between the results for EX, SVEA and MFA calculations. We find good quantitative agreement between the EX and SVEA results and the MFA results show good qualitative agreement only. Hence, the latter can be used for qualitative interpretation of the SVEA and EX numerical results.

The close agreement between our SVEA and EX results is very much dependent upon our choice of the wave vector, (17), to minimize the spacial variation of the envelopes, (15) and (16). This, together with the independent introduction of the background value for the dielectric function into the SVEA, Eq. (20), through the constitutive relation we feel is unique. We know of no other treatment of the SVEA similar to this reported in the literature. It is to be emphasized that the introduction of ϵ_∞ into the SVEA through the constitutive relation is absolutely essential, on physical grounds, to obtain the proper value for the phase velocity appearing in (18). The strong agreement between the SVEA and EX results for steady-state conditions gives credibility to our use of the SVEA, Eqs. (20) - (22a-c), to numerically calculate the switching times, without using adiabatic elimination. Since the excitons strongly couple to the photons and have a large dispersive contribution, the group velocity is significantly changed from the phase velocity and determines the eventual escape time of the photons from the medium.

Furthermore, we have demonstrated in the results shown in Figure 2 that the nonlinearity in excitonic-biexcitonic CuCl which causes OB is of the Kerr type. This suggests further experimental investigations of OB in CuCl for incident laser tuning on either side of the two-photon biexciton resonance.

$$b = e^{-i\omega t}\left[b_F e^{ikx} + b_B e^{-ikx}\right] \quad , \qquad B = e^{-2i\omega t}\left[B_0 + B_F e^{2ikx} + B_B e^{-2ikx}\right] \quad , \tag{21a,b}$$

$$B^+b = e^{i\omega t}\left[B_0^+ b_F e^{ikx} + B_0^+ b_B e^{-ikx} + B_B^+ b_B e^{+ikx} + B_F^+ b_F e^{-ikx} + B_B^+ b_F e^{3ikx} + B_F^+ b_B e^{-3ikx}\right] \quad , \tag{21c}$$

where the last equation follows from the first two.

Using the transformation presented above in Eqs. (7), we arrive at the closed set of equations

$$\frac{\partial b_\nu}{\partial t} = -i(\delta - i\gamma_x)b_\nu + g_1 E_\nu^+ - g_2\left[E_\mu^- B_0 + E_\nu^- B_\nu\right] \quad , \tag{22a}$$

$$\frac{\partial B_0}{\partial t} = -i(\Delta - i\gamma_m)B_0 + g_2\left[E_B^+ b_F + E_F^+ b_B\right] \quad , \qquad \frac{\partial B_\nu}{\partial t} = -i(\Delta - i\gamma_m)B_\nu + g_2 E_\nu^+ b_\nu \quad , \tag{22b,c}$$

where $\nu = F,B$ and μ takes the opposite subscript. Here, $\delta = \omega_x - \omega$ and $\Delta = \omega_m - 2\omega$ and we have neglected triple frequency contributions.

The equations (20) - (22) are appropriate for discussing dynamical behavior and stability conditions[6], as well as the steady state. For our purposes here, however, we focus our attention on the steady-state solutions to these equations.

The steady-state solution for the set of equations (22) gives

$$b_\nu = -\frac{ig_1\Delta' E_F^+}{S}\left[\delta'\Delta' - g_2^2|E^+|^2\right] \quad , \tag{23a}$$

where

$$S = (\delta'\Delta')^2 - 2g_2^2\delta'\Delta'\left(|E_F^+|^2 + |E_B^+|^2\right) + g_2^4\left(|E_F^+|^4 + |E_B^+|^4 + |E_F^+|^2\,|E_B^+|^2\right) \quad . \tag{23b}$$

From (3), (5) and (16), Eqs. (23) can be used to eliminate the polarization terms on the rhs of Eq. (20) in steady state to give,

$$\frac{\partial E_F^+}{\partial x} = \frac{2\pi i\omega\Delta' g_1^2}{cS\sqrt{\varepsilon_r(0)}}\left[\delta'\Delta' - g_2^2|E_F^+|^2\right]E_F^+ - \frac{i\omega\left(\varepsilon_r(0) - \varepsilon_\infty\right)}{2c\sqrt{\varepsilon_r(0)}}E_F^+ \quad , \tag{24a}$$

$$\frac{\partial E_B^+}{\partial x} = -\frac{2\pi i\omega\Delta' g_1^2}{cS\sqrt{\varepsilon_r(0)}}\left[\delta'\Delta' - g_2^2|E_F^+|^2\right]E_B^+ + \frac{i\omega\left(\varepsilon_r(0) - \varepsilon_\infty\right)}{2c\sqrt{\varepsilon_r(0)}}E_B^+ \quad . \tag{24b}$$

It is interesting to note that if the last term in (23b) were a perfect square, Eqs. (24) could then be analytically integrated. However, this is not the case, so we proceed with numerical solution of Eqs. (24) with the appropriate boundary conditions[6].

III. RESULTS AND DISCUSSION

In this discussion, we discuss results of numerical integration of the EX equation, Eq. (11), and the field equations in the SVEA, Eqs. (24) with the appropriate boundary conditions[6]. Comparison is made between these solutions and the corresponding analytical MFA results from Eq. (14).

The EX, SVEA and MFA results for a sample length $L = 10\ \mu m$ are presented in Figure 1. In this case, we fixed the frequency $\omega = 3177$ mev and varied L to obtain a well-defined OB condition in the MFA. This is why $L = 9.986\ \mu m$ in these results. For $L = 1\ \mu m$, no bistability is observed in any of the approximations; this is a strong demonstration of the sensitivity of the OB curves with sample depth. We have chosen a value for ω far from resonance as was also done in reference 4. γ_x is reported to be small[4] and is usually taken to be zero, although it is not entirely negligible. Since both γ_m and γ_x are, in general, affected by impurities and other conditions, we take $\gamma_x = 0$. The values for γ_m vary considerably in the literature; we have taken the value of $\gamma_m = 0.30$ mev here.

where

$$P^{+}(x) = P_F^{+}(x)\, e^{+i\hat{k}x} + P_B^{+}(x)\, e^{-i\hat{k}x} \qquad (16)$$

The corresponding relations for the field are derived from Eq. (3).

At this point the wave vector $\hat{k}$ remains unspecified, whereas the frequency ω is the carrier frequency of the incident laser field. To render the envelope functions, P_F^{+}, P_B^{+}, and E_F^{+}, E_B^{+} as slowly-varying as possible in the medium, we should choose the wave vector $\hat{k}$ to correspond to that in the medium, Eq. (13). Since k in (13) is internal-field-dependent as well as complex, we define $\hat{k}$ by the following,

$$\hat{k} = k_0 \, \mathrm{Re}\, \sqrt{\varepsilon(o)} \qquad (17)$$

where $\varepsilon(o) = \varepsilon(|E| = 0)$.

The second-order Maxwell's equation, Eq. (2), can now be reduced to a first-order differential equation using (15), (16) and the standard conditions[10]. The resulting equations for the forward and backward propagating fields in the medium are,

$$\frac{\partial E_\nu^{+}}{\partial t} + \frac{c^2\hat{k}}{\omega}\frac{\partial E_\nu^{+}}{\partial x} = 2\pi i\omega \left\{ P_\nu^{+} + \frac{2i}{\omega}\frac{\partial P_\nu^{+}}{\partial t} - \frac{[c^2\hat{k}^2 - \omega^2]}{4\pi\omega^2} E_\nu^{+} \right\} , \qquad (18)$$

where we have taken the polarization operators as expectation values, a condition to be assumed from here on and the subscript ν denotes F or B. The minus sign on the left-hand side is taken when $\nu = B$.

It is to be noted that the last two terms on the right-hand side (rhs) of Eq. (18) do not appear in the usual SVEA[10-13]. These terms arise entirely because of our choice of $\hat{k}$. In the usual SVEA, k is chosen to be equal to ω/c. Since the nonlinearity is large in excitonic CuCl, the usual procedure, i.e., choosing $k = k_0$, would mean the variables in fact would not be so slowly-varying as to meet the criteria of the SVEA[10-13].

The second terms in the brackets on the rhs of Eq. (18) make contributions on the order of the first terms on the left-hand side (lhs) in these equations via the constitutive relations. Thus, we can replace the time derivative on the rhs by the supplementary constitutive relation

$$\frac{\partial P_{F,B}^{+}}{\partial t} = \frac{[\varepsilon_r(o)-1]}{4\pi}\frac{\partial E_{F,B}^{+}}{\partial t} , \qquad (19)$$

where $\varepsilon_r(o) = \mathrm{Re}\, \varepsilon(o)$.

The contribution from the time derivative of the higher-order terms in ε which have been dropped from (19) are entirely negligible for CuCl. The constitutive relation (19) is entirely consistent with our choice, Eq. (17). It is to be noted here that $\varepsilon(o)$ in (15) and (17) contains ε_∞ by Eq. (9) which was injected phenomenologically into the model calculation and is essential for a meaningful constitutive relation for excitonic CuCl[1-6].

If Eq. (15) is combined with Eq. (18), we have, using the notation in Eq. (18),

$$\frac{\partial E_\nu^{+}}{\partial t} \pm \frac{c}{\mathrm{Re}\sqrt{\varepsilon_r(o)}}\frac{\partial E_\nu^{+}}{\partial x} = \frac{2\pi i\omega}{\varepsilon_r(o)} \left\{ P_\nu^{+} - \frac{[c^2\hat{k}^2 - \omega^2\varepsilon_\infty]}{4\pi\omega^2} E_\nu^{+} \right\} . \qquad (20)$$

The factor onto the spatial derivatives on the lhs of (20) is just the phase velocity of the field envelope in the medium. We have not included the small imaginary contribution to the dielectric constant in the definition of the envelope.

Equation (20) can be solved with the appropriate boundary and initial conditions when P_F^{+} and P_B^{+} are related to the respective fields by the full constitutive relations determined from the equations of motion obtained from the Hamiltonian (1) and the relations (6).

The specification of the appropriate boundary conditions consistent with our model, to be used for the solutions of Eq. (20) are discussed elsewhere[6] and will not be presented here.

In order to integrate Eq. (20) with the appropriate boundary conditions[6], we must determine the polarizations P_F^{+} and P_B^{+} from the material equations of motion. If we use Eqs. (15) and (16) in (5), then we have the identifications

frequency of the incident electric field, where we have used

$$E^{+}(x,t) = e^{-i\omega t}E^{+}(x) \quad , \tag{10}$$

where x is the longitudinal direction of propagation in the dielectric medium of length L.

The stationary solution of (2) based upon (9) and (10) is

$$\frac{\partial^2 E(x)}{\partial x^2} + \varepsilon k_0^2 E(x) = 0 \tag{11}$$

where $k_0 = \omega/c$ is the wave number in vacuum, and we have dropped the superscripts on E and ε for convenience. The details of the solution of (11) with the appropriate boundary conditions in terms of the incident intensity, I_{IN} as a function of the transmitted intensity I_0 is given elsewhere and will not be reported here.

B. Mean-Field Approximation (MFA) Solution

The MFA often used in the literature to describe OB in a Kerr medium in a Fabry-Perot cavity avoids the necessity for numerical integration by imposing the ansatz,

$$E(x) = E_R(0)\, e^{ikx} + E_L(0)\, e^{-ikx} \quad , \tag{12}$$

where $E_R(0)$ and $E_L(0)$ are rightward and leftward propagating components of the electric field E and are taken as independent of x. Here k is the complex wave vector defined by

$$k = k_0 \sqrt{\varepsilon} \quad , \tag{13}$$

where ε is given by [illegible]

[illegible] with the boundary conditions, gives[5]

$$\left| \frac{E_R(0)}{E_L(0)} \right|^2 e^{i(\varphi_R - \varphi_L)}$$

$$\left(\frac{\quad}{\quad} \right)^{1/2} \tag{14}$$

[illegible] the nonlinear coefficient of the CuCl material. Equation (14) [illegible] cannot be regarded as reliable under arbitrary [illegible] numerical results and a qualitative understanding [illegible] conditions by studying this simple result.

[illegible] using the EX and SVEA that the resulting [illegible] field-dependent resonance condition suggested [illegible] normally occurs far removed from any [illegible]

[illegible]

The [illegible] theory and is a well-known procedure[10-12] which [illegible] where the material amplitudes are slowly-[illegible] the [illegible] of the second-order Maxwell's equation [illegible] the [illegible] removing the rapidly-varying part from the [illegible] carrier frequency), resulting in a [illegible] Numerical integration using the SVEA can correspond to a drastic [illegible] integration of the corresponding second-order Maxwell equation with associated boundary and initial conditions. Since the resulting Maxwell's equation in the SVEA retains first-order spatial and time derivatives, it is expected that the results using the SVEA will be generally in closer agreement with the EX results than those using the MFA.

As stated earlier in this section, the operators and field variables in (3) are understood to depend upon the time t and propagation distance x. It is thus useful to transform the relations (3) to slowly-varying forward E and backward B propagating components. For the polarization P, Eq. (3), we define the relation

$$P = e^{i(kx-\omega t)}P(x,t) + e^{-i(kx-\omega t)}P^{\dagger}(x,t) \quad , \tag{15}$$

with positive and negative frequency components, E^+ and E^-, respectively, is assumed to be the superposition of rightward and leftward propagating monochromatic plane waves in the dielectric medium and is understood to have both spacial and time dependence in general. The exciton and biexciton operators are correspondingly understood to have k-vector dependence consistent with the electric field. We take the electric field E to be classical and proceed with the semiclassical model calculation where the field and dielectric are coupled by Maxwell's equation

$$c^2 \frac{\partial^2 E}{\partial x^2} - \frac{\partial^2 E}{\partial t^2} = 4\pi \frac{\partial^2 P}{\partial t^2} \quad , \tag{2}$$

where P is the polarization in the material. In terms of positive and negative frequency components,

$$E = E^+ + E^- \quad , \qquad P = P^+ + P^- \quad , \tag{3}$$

and consistent with the rotating wave approximation

$$H' = -E^+P^- - E^-P^+ \quad . \tag{4}$$

Thus, from (4) and (1),

$$P^+ = ig_1 b + ig_2 b^+ B \, . \tag{5}$$

From this relation we obtain the dielectric function[3] $\varepsilon^{\pm}$,

$$\varepsilon^+ = 1 + 4\pi \langle P^+ \rangle / E^+ \quad . \tag{6}$$

To obtain the explicit stationary expression for (6) requires solution of the equations of motion from (1) in the steady state. The equations of motion are

$$i \frac{\partial b}{\partial t} = \omega_x b + ig_1 E^+ - ig_2 E^- B - i\gamma_x b \quad , \tag{7a}$$

$$i \frac{\partial B}{\partial t} = \omega_m B + ig_2 E^+ b - i\gamma_m B \quad , \tag{7b}$$

$$i \frac{\partial b^+ B}{\partial t} = -(\omega_x - \omega_m) b^+ B + ig_1 E^- B + ig_2 E^+ (b^+ b - B^+ B) - i\gamma b^+ B \quad , \tag{7c}$$

where we have added phenomenological relaxation rates γ_x, γ_m and γ for the excitons, biexcitons and field-exciton-biexciton interaction (FEB), respectively. This hierarchy would close with the addition of a decoupling of the atom-field moments and a fourth equation for $(b^+b - B^+B)$. However, this expression for the difference of the populations appears only in (7c), and the contribution of (7c) itself in steady state only adds a small correction[1] to the nonlinear dielectric function determined from (5). Therefore, we set

$$\langle b^+ b - B^+ B \rangle = S_0 \quad , \tag{8}$$

its thermodynamic equilibrium expectation value.

A. Exact Solution (EX)

We proceed by solving (7a) and (7b) together in steady state to obtain the main contribution to the polarization in Eq. (5), i.e., we ignore the small FEB correction from the second term in (5) which is consistent with termination of the hierarchy with (7a) and (7b)[3]. This, together with Eq. (6), yields

$$\varepsilon^+ = \varepsilon_\infty + 4\pi \left[\frac{g_1^2}{\Delta' - \frac{g_2^2 |E|^2}{\Delta^*}} \right] \quad , \tag{9}$$

where we have replaced unity in the first term in (9) by the high frequency dielectric constant ε_∞. ε_∞ represents the contribution of high frequency modes in the medium not accounted for in our Hamiltonian. Here $\Delta' = \omega_x - \omega - i\gamma_x$, $\Delta^* = \omega_m - 2\omega - i\gamma_m$, and $|E|^2 \equiv E^+E^-$. In these relations ω is the

STATIONARY PROPERTIES AND SWITCHING CHARACTERISTICS OF DISPERSIVE OPTICAL BISTABILITY IN $CuCl$

C. M. Bowden and J. W. Haus*
Research Directorate, US Army Missile Laboratory
US Army Missile Command
Redstone Arsenal, Alabama 35898, USA

and

C. C. Sung
Department of Physics
The University of Alabama in Huntsville
Huntsville, Alabama 35899, USA

SUMMARY

The stationary properties are presented for optical bistability (OB) from a semiclassical exciton-biexciton model for $CuCl$ in the limit of large Fresnel number. Numerical solutions using the second-order Maxwell's equation (EX) are compared with corresponding solutions where the slowly-varying envelope approximation (SVEA) has been made. In addition, the EX and SVEA results are compared with the corresponding mean-field results (MFA). It is found that results obtained using the SVEA are in close quantitative agreement with those using the second-order Maxwell's equation (EX results), whereas the MFA results are in qualitative agreement only. On the basis of the close quantitative agreement between the SVEA and EX results, we use the SVEA hierarchy of equations to numerically compute dynamics of switching as the input field is varied. The calculation is done without adiabatic elimination, which is not valid for OB in $CuCl$. Our main conclusions are that the OB in $CuCl$ in steady state can be characterized very closely as a Kerr type nonlinearity and that the switch-down time is approximately equal to the switch-up time which is close to 80 psec at the laser photon energy 3177 mev. The switching time correlates with the polariton lifetime in the material and this has significant consequences for the switching characteristics in dispersive media.

I. INTRODUCTION

The subject of this paper is the analysis of steady-state characteristics and dynamics of optical bistability using the excitonic nonlinearity in $CuCl$[1-6]. Here, the physics of the excitonic and biexcitonic contributions to the nonlinear index of refraction is understood and a simple model, which does not include the details of the band structure, should suffice to describe quantitatively the salient features of OB in this system. Some preliminary experimental results have demonstrated the existence of OB[7,8]. In this work, we improve considerably, with respect to previous works, on the calculations of the output intensity as a function of input field in steady state. We emphasize at the outset that our calculations are not intended to fit any specific experimental data. Since there are inconsistencies between some published results[1,7,8], we vary our parameters to compare our results with other published works.

Previous calculations on $CuCl$ assumed the dielectric function to be constant throughout the cavity; this is a version of the so-called "mean-field approximation" (MFA), and has the advantage of yielding analytical results[1-3,5,7]. Since the limit is idealized, we solve here, in addition, the second-order Maxwell equation which requires numerical integration. This we refer to as "exact" (EX) results. In addition, we establish the quantitative credibility of the slowly-varying envelope approximation (SVEA) for our model, and use this to numerically calculate the dynamics of switching and switching times for $CuCl$ without adiabatic elimination of the excitonic and biexcitonic variables, which elimination procedure is invalid in this case[9].

The model for excitonic-biexcitonic $CuCl$ which we use[3] and which has been utilized by others[4], is presented in the next section. In that section we develop the framework for computing the EX results from the model and the boundary conditions. Next, we define the MFA as applied to dispersive OB for this model. Finally, the equations of motion are presented in the SVEA, which are used to compute the switching dynamics and from which the steady-state relations are developed. Section III is used to discuss the results and comparisons for the EX, SVEA, and MFA; and to present the results for the dynamics and the switching times. Our conclusions are presented in the final section.

II. MODEL HAMILTONIAN AND STATIONARY SOLUTION

The model Hamiltonian[1-6] for $CuCl$ which is useful for calculating the nonlinear interaction with light for an incident laser field tuned near the exciton or two-photon biexciton resonance, is given in the rotating wave and electric dipole approximation, by $H = H_0 + H'$, where

$$H_0 = \omega_x b^+ b + \omega_m B^+ B \quad , \tag{1a}$$

$$H' = ig_1 E^+ b^+ + ig_2 E^+ B^+ b + h.c. \tag{1b}$$

Here, $B^+(B)$ and $b^+(b)$ are the creation (annihilation) operators for the biexcitons and excitons, respectively, and ω_m and ω_x are the respective transition energies (units such that $\hbar = 1$ are used); g_1 and g_2 are the coupling constants whose numerical values are inferred from experiments. The electric field E

*A National Research Council Research Associate

This condition corresponds to $\bar{n}_c$ approximately at the mid-point of the unstable region of semiclassical, deterministic bistability. This is depicted in Figure 3 which exhibits the analytically determined deterministic semiclassical result for two, two-level atoms[10]. Also indicated in the figure, for reference, is the approximate region of the heuristic study.

Figure 4 shows the emergence of the bimodal dependence of the inversion, Eq. (20), for seven atoms as the detector response time T_f is increased from $T_f = 0$, Figure 4a, to $T_f = 1$, Figure 4b. Bimodality emerges when the response time is greater than or approximately equal to the atomic decay time. The bimodality appears when $T_f \geq 1$ due to the filtering of the Rabi cycling by the detector response.

A histogram of the up times, t_u, and down times, t_d, sampling during the time interval of Figure 4 is shown in Figure 5 for four different values for T_f. The average up time, T_u, and down time, T_d, corresponding to each value for T_f as indicated in the figure, is seen to increase markedly as T_f approaches unity, consistent with the approach to bimodality shown in Figure 4.

The inversion, Eq. (20) as a function of the incident field E for the conditions for $T_f = 1$ and otherwise the same as those of Figures 4 and 5 is shown in Figure 6 for one cycle. The bimodality of the inversion for a fixed field condition exhibited in Figure 4 is manifest as a hysteresis in Figure 6 as the field E is varied over one cycle. Further cycling, of course, leads to variations in the hysteresis profile due to statistical fluctuations. Successive cyclings then comprise an ensemble of measurements with corresponding statistical variations.

IV. CONCLUSIONS

We have shown that the LFC applied to the semiclassical, macroscopic Maxwell-Bloch formulation causes OB in steady-state in the absence of any resonator or optical feedback conditions[10]. The hysteresis is absorbance-dominated as shown in Figures 1 and 2, i.e., the overall reflectivity is negligible, which is sufficient evidence that the OB is independent of any optical feedback. This is consistent with the fact that the thickness of the macroscopic material is much smaller than a wavelength of the illuminating laser. This was necessary, since the density requirement for OB corresponds to an optically opaque material. Thus, the propagation aspect of the problem is actually an optical "skin effect".

Furthermore, we have shown in Section III that OB occurs in a microscopic model of laser-driven, two-level atoms, also without any cavity or optical feedback conditions. This gives credence to the validity of using the LFC in the macroscopic formulation when the density of material is sufficiently high, i.e., near that for solids. Furthermore, we have shown that the microscopic model gives OB when fluctuations are included, and the results are given in Figures 3 - 6.

These results should give impetus to stimulate further investigations into a wide range of materials both theoretically and experimentally, to establish the kind of intrinsic OB discussed here. We feel that further investigations of this type are of profound fundamental interest in relating microscopic electrodynamics to the macroscopic and the origin of the local field effect as well as having obvious important practical implications.

REFERENCES

1. C. M. Bowden and C. C. Sung, Phys. Rev. A19, 2392 (1979).

2. C. M. Bowden, "Cooperative Optical Bistability in a Small Volume Without Mirrors", IEEE Cat. No. 80 CH 1561-0, XI IQEC, Boston, MA, June 23-26, 1980, p. 589.

3. C. M. Bowden, "Optical Bistability Based Upon Atomic Correlation in a Small Volume", in Optical Bistability, eds., C. M. Bowden, M. Ciftan and H. R. Robl (Plenum, NY, 1981), p. 405.

4. C. Flytzanis and C. S. Tang, Phys. Rev. Lett. 45, 441 (1980).

5. J. Hajto and J. Janossy, Phil. Mag. B47, 347 (1983).

6. D. A. B. Miller, A. C. Gossard and W. Wiegmann, Opt. Lett. 9, 162 (1984).

7. M. Dagenais and W. F. Sharfin, "Cavityless Optical Bistability due to Light-induced Absorption in Cadmium Sulfide", Applied Physics Letters, to be published.

8. K. Henneberger and H. Rossman, Phys. Stat. Sol.(b) 121, 685 (1984).

9. J. Van Kronendonk and J. E. Sipe, in Progress in Optics XV, edited by E. Wolf (North-Holland, Amsterdam, 1977), p. 245.

10. F. A. Hopf, C. M. Bowden and W. H. Louisell, Phys. Rev. A29, 2591 (1984); in Optical Bistability II, eds., C. M. Bowden, H. M. Gibbs and S. L. McCall (Plenum, NY, 1984), p. 361.

11. See papers and references in: Optical Bistability, eds., C. M. Bowden, M. Ciftan and H. R. Robl, (Plenum, NY, 1981); Optical Bistability II, eds., C. M. Bowden, H. M. Gibbs, and S. L. McCall (Plenum, NY, 1984).

12. J. D. Jackson, "Classical Electrodynamics", 2nd Edition, (Wiley, NY, 1962), Chapter 4.

13. A. Icsevgi and W. E. Lamb, Phys. Rev. 185, 517 (1969).

14. L. M. Narducci, S. S. Mitra, R. A. Shatas and C. A. Coulter, Phys. Rev. A16, 247 (1977).

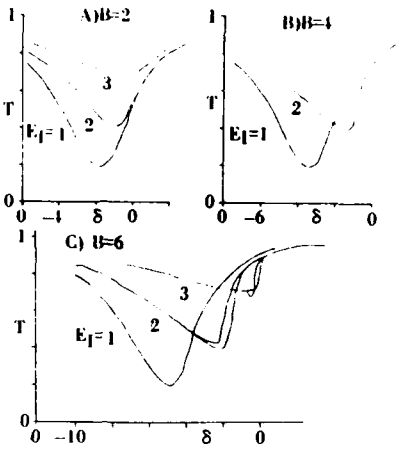


Figure 1. Transmission T vs. detuning δ for (a): B = 2 and ε_I = 1,2,3. (b) B = 4 and ε_I = 1,2. (c) B = 6 and ε_I = 1,2,3.

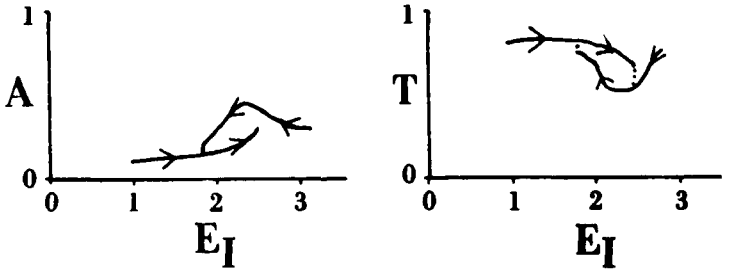


Figure 2. (a) Absorbance A. (b) Transmission T vs. input field ε_I for Δ = -1.5 and B = 6.

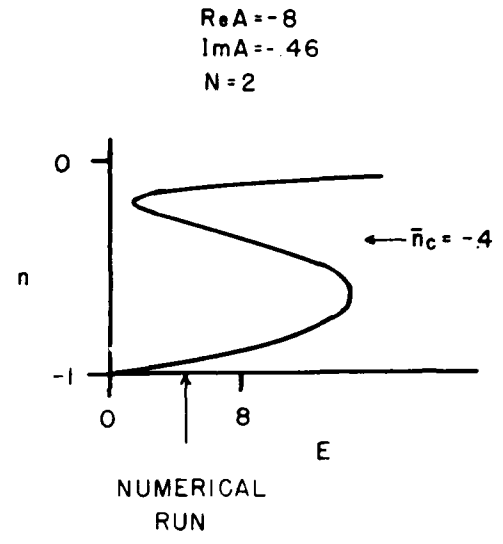


Figure 3. Deterministic steady-state for the inversion n vs. incident field of amplitude E for two, two-level atoms for Re A = -8, Im A = -0.46. The corresponding position of the parameter $\bar{n}_c$ used in the stochastic model is shown together with the E-value and approximate location in the hysteresis zone where the stochastic numerical calculations were carried out indicating bimodal dependence.

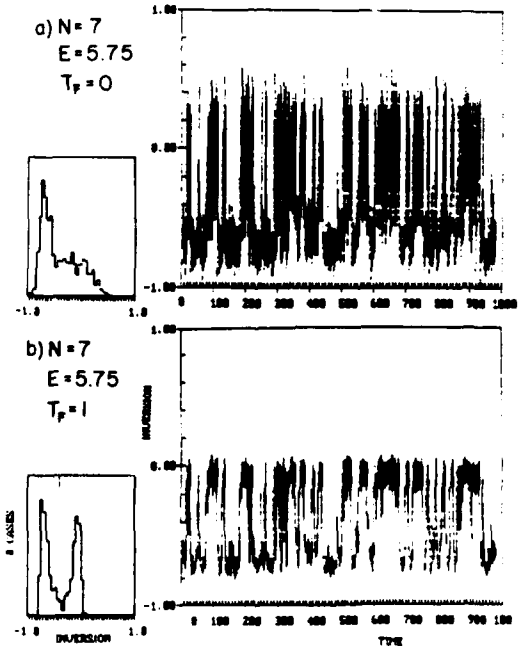


Figure 4. (a) Inversion $\bar{n}(t)$ as a function of time for N = 7, E = 5.75 and T_F = 0. Inset is the histogram of the number of occurrences of values of the inversion $\bar{n}(t)$ in the time interval. (b) Same as for (a), except that T_F = 1.

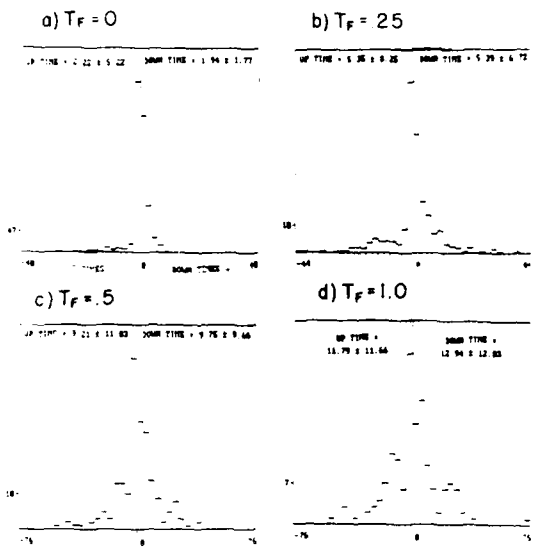


Figure 5. Histograms of the occurrences of the up, t_u, and down, t_d, times in the time interval of Figure 4. (a) $T_f = 0$. (b) $T_f = 0.25$. (c) $T_f = 0.5$. (d) $T_f = 1.0$. The average up (down) times $T_u(T_d)$ are given for each case in the figure.

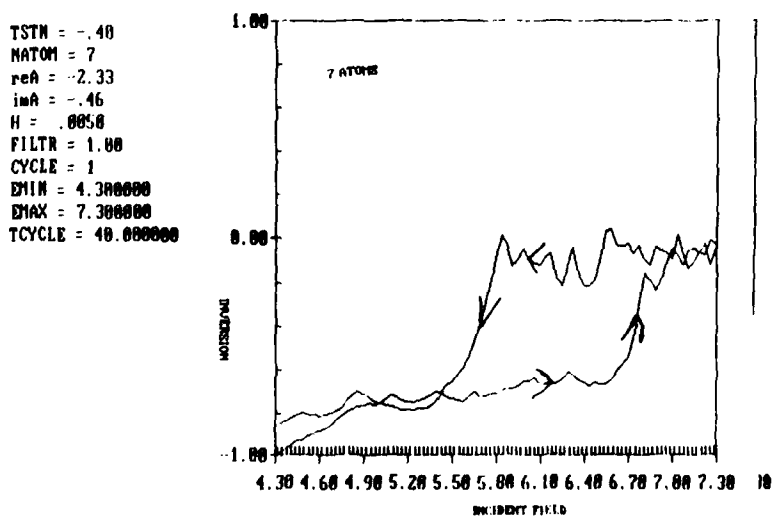


Figure 6. Hysteresis cycle of the inversion $\tilde{n}(t)$ vs. incident field E for N = 7, $\tilde{n}_c$ = -0.4, Re A = - 2.33, Im A = -0.46, T_f = 1.0.

PROSPECTS FOR PARALLEL NONLINEAR OPTICAL SIGNAL PROCESSING USING GaAs ETALONS AND ZnS INTERFERENCE FILTERS

by
H. M. Gibbs, J. L. Jewell,* Y. H. Lee, A. Macleod, G. Olbright,
S. Ovadia, N. Peyghambarian, M. C. Rushford,
M. Warren, and D. A. Weinberger

Optical Sciences Center
University of Arizona
Tucson, Arizona 85721

*Present address: AT&T Bell Labs, Holmdel, N. J. 07733

and

T. Venkatesan
Bell Communications Research
Murray Hill, New Jersey 07974

ABSTRACT

Micron-thick semiconductor etalons can perform logic operations using input beams of a few mW. Problems associated with the parallel operation of thousands of etalons are discussed. GaAs and ZnS appear to be the most attractive materials for low-power room-temperature operation.

INTRODUCTION

Semiconductors are attractive for all-optical logic because of their strong interactions with light: 50% of an incident beam can be absorbed in a few μm or less. If a Fabry-Perot interferometer is constructed containing a semiconductor, the transmitted intensity I_T versus the input intensity I_I can exhibit any of the desired logical operations: optical transistor (AC gain), optical memory (optical bistability), optical limiting, AND gate, NOR gate, etc.[1,2]

Single-beam or two-beam activation of a nonlinear etalon could be used for very-high-speed (≈1 ps) switching for multiplexing and encryption. Nonresonant nonlinearities can be selected with very fast recovery times permitting many-GHz operating frequencies. Resonant band-edge nonlinearities in semiconductors usually have recovery times governed by carrier lifetimes, typically several ns. Techniques for reducing the carrier lifetimes are being applied to bistable etalons in an attempt to achieve GHz-plus frequencies, i.e., sub-ns times between ps decisions.

For parallel computations, the nonlinear etalons are already interestingly fast. For example, 10^6 operations in parallel at a kHz rate is equivalent to single-channel operation at a GHz rate. Progress toward parallel operation using GaAs and ZnS etalons is emphasized here. Nonlinear etalons should be useful as spatial light modulators and as decision-plane discriminators in correlators.

GaAs

a) Bulk versus superlattice

Our GaAs etalons are grown by molecular beam epitaxy and usually contain 1 to 4 μm of GaAs. Sometimes bulk GaAs is used and sometimes a superlattice of alternating layers of GaAs and $Al_{0.3}Ga_{0.7}As$ with layer thicknesses ranging from 50 to 350 Å. At room temperature, the free-exciton resonance is more pronounced in the superlattice, but the optical bistability results are quite similar (same minimum switch-on power, similar wide loops at high powers, etc.). Admittedly, this similarity suggests that the mechanism for bistability may not be excitonic. However, bistability is seen with ≤ 3 kW/cm², which calculations show is sufficient to give a large enough refractive index change for bistability if the mechanism is excitonic but not if it is band filling. The explanation for the similarity is under study, but may be as follows. Background absorption α_B (from the band tail which is filled or saturated only at considerably higher intensities) prevents on-resonance purely absorptive bistability. Dispersive bistability can only occur for large enough detunings so that α_B is small enough. Operation may be forced even farther off resonance by the fact that the exciton absorption does not decrease at all frequencies as the intensity increases; the absorption first increases and then decreases as the intensity is increased for $\Delta \approx 1$ to 3 [$\Delta \equiv (\nu_{Ex}-\nu)/\delta\nu_{Ex}$ where ν_{Ex} is the frequency of peak exciton absorption, ν is the laser frequency, and $\delta\nu_{Ex}$ is the exciton's half-width at half-maximum]. With $\Delta > 3$, the exciton absorption has to be very nearly saturated to obtain an index change adequate for bistability. Before switch-on is reached, the superlattice exciton absorption profile is broadened and its peak reduced, so that it appears much like the bulk profile. Then for higher intensities, the bulk and superlattice etalons behave similarly. Perhaps a thorough understanding of this similarity will lead to lower-power operation using superlattices; meanwhile, the only incentive for using a superlattice is the ability to tune the exciton wavelength by varying the GaAs well thickness. This has permitted room-temperature operation of a bistable etalon using a diode laser.[3]

b) Parallel operation

For parallel operation of many beams on the same etalon, individual pixels may be defined by the focal spots themselves, or by etching, or by proton damage. Several factors must be considered in designing etalons for parallel operation: crosstalk via diffraction, diffusion, and luminescence; thermal limitations; uniformity in thickness; etc.

Numerical simulations in one transverse dimension[4] and in two transverse dimensions[5] show that beams separated by three or more spot diameters function without crosstalk; i.e., if all of the surrounding spots are switched "on" simultaneously, the spot in question remains "off." Those calculations assumed a defocusing nonlinearity and a Fresnel number of order one to assure whole-beam switching. Thus crosstalk via light diffraction would permit arrays of 3-μm spots separated by 10 μm, for example.

Experiments in GaAs etalons illustrate that carrier diffusion can impose larger separations. When lasing was observed[6] in 4-μm-thick GaAs etalons, some critics questioned that carriers excited at the entrance face would diffuse across the sample. However, recent measurements show that carriers can diffuse tens of μm.[7] If a pixel separation of less that 100 μm is desired, it can be achieved in at least two ways. One way is to define individual pixels by removing material between pixels by plasma etching or by proton damaging the same material. In the latter case, the carriers would recombine in the damaged inter-pixel material. A second method is to reduce the carrier lifetime within the pixel, possibly by removing the top AlGaAs window leading to fast surface recombination or by proton damaging the entire etalon leading to fast bulk recombination. Research on reducing the carrier lifetime without increasing the operating powers is underway; if sucessful, it will remove the problem of crosstalk by diffusion as well. Hope for success is based on the fact that the exciton lifetime at room temperature is subpicosecond,[8] so reduction of the carrier lifetime from several ns to 100 ps or less need not necessarily reduce the exciton lifetime.

Luminescence may also result in crosstalk between pixels. In an array defined by etching, light diffraction perpendicular to the plane and the solid angle factor may make the crosstalk acceptably low. If not, the inter-pixel space can be filled with an absorbing material.

Thermal problems do not have solutions as simple and straightforward as those for crosstalk problems. Both bulk and superlattice etalons have shown optical bistability with less than 10 mW cw incident. Crosstalk considerations above suggest that 10-μm separations are reasonable, resulting in 10^6 pixels per cm^2. If each pixel is illuminated by 10 mW and most of that light is absorbed in the pixel, then 10 kW/cm^2 must be removed from the crystal. Heat removal of 100 W/cm^2 is common for electronics; 1 kW/cm^2 is conceivable, but 10 kW/cm^2 is highly unlikely. Thus the separation must be increased or the duty factor reduced to 0.1 or 0.01. For many applications, cw operation should not be needed. If the etalon is used as an array of NOR gates,[9] for example, a logic decision can be made in 1 ps.[10] The device is then allowed to recover in total darkness; this requires 5 to 10 ns. If each logic operation can be performed using 1 to 10 pJ per pixel, 10^6 pixels/cm^2 can be operated at 100 MHz to 10 MHz to keep the thermal load at 100 W/cm^2. Recently we observed NOR gating in a GaAs superlattice (76Å wells) with 5 to 1 contrast and ≈40 pJ absorbed in a ≈20 μm spot.[11] Hopefully device improvements will result in lower energy dissipation. A 10-MHz rate is, of course, exceedingly fast compared with the 10 to 100 Hz operation of the spatial light modulators used in most current parallel optical processors.

Uniformity of thickness is needed so that the etalon's peak transmission frequency is independent of position on the etalon. In the first GaAs devices a "very flat" device always had at least a one-order thickness variation across a 1-mm-diameter-etched region. By using two extra stop layers, devices were constructed with a small fraction of an order of variation.[12] Reactive plasma etching promises even flatter devices.[13] So far, flatness has been limited by etching techniques rather than by growth imperfections. Plasma etching is also attractive because it can produce almost vertical walls. Already GaAs arrays have been etched on samples unsuitable for bistability: 5×5 μm^2 pixels separated by 10 μm and 1-μm-diameter cylindrical pixels with 5-μm spacing.

ZnS AND ZnSe INTERFERENCE FILTERS

Karpushko and Sinitsyn[14,15] were apparently the first to observe passive optical bistability in a semiconductor using nonlinearities of the ZnS intermediate layer of a dielectric interference filter. The shortest conceivable optical resonator would have a length of $\lambda/2n_0$. These devices are almost that short, having a nonlinear layer only λ/n_0 thick, i.e., 0.22 μm for their 514.5-nm filter. Using ZnS, they have seen 10-μs switching times and 4kW/cm^2 switch-on intensities. At first we failed to reproduce those values;[16] our commercial 514.5-nm interference filter showed msec switching times and required about 6 kW/cm^2. Also a "damage" effect shifted the peak of the etalon's transmission, presumably because water vapor was driven from the filter by heating; the vapor and peak returned gradually in several months. This damage, which may have been associated with glue holding on a protective glass cover, is not seen in new unprotected filters grown by the Thin Films Group at the University of Arizona. Furthermore <10 μs switching times and 50 kW/cm^2 intensities have been seen in the new filters.

There is still uncertainty about the origin of the optical nonlinearity. Karpushko and Sinitsyn[15] attribute it to two-photon photorefraction which they state occurs in thin-film but not bulk material. Our observations show that both the sign (positive) and response times of the nonlinearity are consistent with a thermal effect. The fact that bistability has been observed at 468, 514.5, and 632.8 nm using ZnS filters points to a nonresonant mechanism. Preliminary results yield faster responses and lower powers the shorter the wavelength. Perhaps further research will lead to lower powers and faster times.

Regardless of origin of the nonlinearity, the ZnS interference filters are attractive for parallel operation. The thin films are produced by evaporation, a standard and relatively inexpensive technique resulting in only a few-nm variation in the peak wavelength across a 5-cm-square filter. An array of 10^6 spots with 10 mW per spot could run cw with no more than 100 W/cm^2 heat load on such a filter assuming 25% absorption. Of course, that would require 10 kW of optical power (at the etalon) and 2.5 kW of heat to be removed.

SUMMARY

GaAs etalons and ZnS interference filters are both attractive for multiple-beam nonlinear optical signal processing. Clearly it is highly desirable to reduce the switching power and heat generated in each pixel. Mea while the present values are low enough to pursue the other problems of generating multiple beams and developing useful architectures.

REFERENCES

1. H. M. Gibbs, S. L. McCall, and T. N. C. Venkatesan, "Optical bistability," Optics News **5**, 6 (1979).

2. H. M. Gibbs, S. L. McCall, and T. N. C. Venkatesan, "Optical bistable devices: the basic components of all-optical systems?" Opt. Engineering **19**, 463 (1980).

3. S. S. Tarng, H. M. Gibbs, J. L. Jewell, N. Peyghambarian, A. C. Gossard, T. Venkatesan, and W. Wiegmann, "Use of a diode laser to observe room-temperature, low-power optical bistability in a GaAs-AlGaAs etalon," Appl. Phys. Lett. **44**, 360 (1984).

4. K. Tai, J. V. Moloney, and H. M. Gibbs, "Optical crosstalk between nearby optical bistable devices on the same etalon," Opt. Lett. **7**, 429 (1982).

5. K. Tai, H. M. Gibbs, J. V. Moloney, D. A. Weinberger, S. S. Tarng, J. L. Jewell, A. C. Gossard, and W. Wiegmann, "Self-defocusing and optical crosstalk in a bistable optical etalon," p. 415 in C. M. Bowden, H. M. Gibbs, and S. L. McCall, eds., Optical Bistability ? (Plenum Press, NY, 1984).

6. A. Passner, H. M. Gibbs, A. C. Gossard, S. L. McCall, T. N. C. Venkatesan, and W. Wiegmann, "Ultrashort laser: lasing in MBE GaAs layer with perpendicular-to-film optical excitation and emission," IEEE J. Quantum Electron. **QE-17**, 1283 (1980).

7. A. Olsson, D. J. Erskine, Z. Y. Xu, A. Schremer, and C. L. Tang, "Nonlinear luminescence and time-resolved diffusion profiles of photoexcited carriers in semiconductors," Appl. Phys. Lett. **41**, 659 (1982).

8. D. A. B. Miller, D. S. Chemla, D. J. Eilenberger, P. W. Smith, A. C. Gossard, and W. T. Tsang, "Large room-temperature optical nonlinearity in GaAs/$Ga_{1-x}Al_xAs$ multiple quantum well structures," Appl. Phys. Lett. **41**, 679 (1982).

9. J. L. Jewell, M. C. Rushford, and H. M. Gibbs, "Use of a single nonlinear Fabry-Perot etalon as optical logic gates," Appl. Phys. Lett. **44**, 172 (1984).

10. A. Migus, A. Antonetti, D. Hulin, A. Mysyrowicz, H. M. Gibbs, N. Peyghambarian, and J. L. Jewell, "One-picosecond optical NOR gate at room temperature with a GaAs-AlGaAs multiple-quantum-well nonlinear Fabry-Perot etalon," submitted to Appl. Phys. Lett.

11. J. L. Jewell, Y. H. Lee, M. Warren, M. C. Rushford, H. M. Gibbs, N. Peyghambarian, A. C. Gossard, and W. Wiegmann, "Optical logic in GaAs Fabry-Perot etalons," submitted to International Commission for Optics - 13, "Optics in Modern Science and Technology," Sapporo, Japan, Aug. 20-24, 1984.

12. J. L. Jewell, H. M. Gibbs, A. C. Gossard, A. Passner, and W. Wiegmann, "Fabrication of GaAs bistable optic devices," Mater. Lett. **1**, 143 (1983).

13. D. Bollinger, S. Iida, and D. Matsumoto, "Reactive ion etching: its basis and future," Solid State Technology, p. 111, May 1984.

14. F. V. Karpushko and G. V. Sinitsyn, "An optical logic element for integrated optics in a nonlinear semiconductor interferometer," J. Appl. Spectrosc. USSR **29**, 1323 (1978).

15. F. V. Karpushko and G. V. Sinitsyn, "The anomalous nonlinearity and optical bistability in thin-film interference structures," Appl. Phys. B. **28**, 137 (1982).

16. D. A. Weinberger, H. M. Gibbs, C. F. Li, and M. C. Rushford, "Room-temperature optical bistability in thin-film interference filters," J. Opt. Soc. Am. **72**, 1769 (1982).

ACKNOWLEDGEMENTS

We would like to extended our thanks to A. C. Gossard and W. Wiegmann of AT&T Bell Labs, Murray Hill, New Jersey, for the growth of the GaAs samples. The Arizona portion of this work has been supported by the U. S. Air Force Office of Scientific Research, U. S. Army Research Office, and the National Science Foundation.

DISCUSSION

S.D.Smith, UK

Could you clarify the latest state of intensity required for the GaAs bistability and nonlinearity? David Miller said at one point that he was getting a saturation of the exciton at a much lower intensity (like 10 or 100 times lower) than the intensity at which bistability was being observed, thus putting in some doubt whether the exciton really had anything to do with it at all.

Author's Reply

The bistability intensities, when we use large beam diameters in order to avoid carrier diffusion, are of the order of 1 to 3 kilowatts per square centimetre. The kinds of saturation intensities that he has measured and that we actually measured at low temperature are of the order of 100 or 150 watts per sq. centimetre. These are somewhat different, but there are a couple of things to remember. The saturation data were taken on-resonance; we are forced to operate off-resonance. The other feature is that when one deduces as low as 100 watts per sq. centimetre one does an analysis to conclude what the plane wave value would be. In fact, people use Gaussian beams; if you look at this data, for example, and just ask the simple question: here is absorption and it goes to a very low value; at what intensity does it fall to one half of its value? It's of the order of a kilowatt per sq. centimetre. Basically if we take his data it's perfectly consistent with ours when we take into consideration the Gaussian profile and the fact that we operate off resonance. And so we believe that the mechanism is excitonic, not that we care a great deal in the sense that it works, and we'll take whatever it is; but one would like to understand why one can't do better in the sense that if one could get rid of background absorption and come in closer. So we are trying to make that case tight and convincing in the sense of doing a careful job of modelling of what our system really looks like, and then one can always hope that the model will lead to improvement if one understands what the limitations are.

A.C.Walker, UK

Can we ask the question about the interference filters – we pointed out that we did not appear to see strong nonlinear effects in purely zinc sulphide systems. You are seeing them; have you tried doing the reverse as we did and that is look at a ZnSe device?

Author's Reply

Yes, we've looked at ZnSe; we see bistability, and it is not very different from ZnS; but it takes a little bit higher powers and the response times were a little bit slower at the wavelengths (632.8 nm) where we've operated. There are no dramatic differences. I do not understand why you would not see it in zinc sulphide.

A.C.Walker, UK

We probably would if we went to higher intensities. I would say our experience is the reverse: that for some reason the ZnSe seems to be the more nonlinear medium at this wavelength (514.5 nm).

Author's Reply

For ZnSe we've used a 632.8 nm interference filter with 3.7 nm FWHM which showed 50 μs switch-on and 75 μs switch-off times with 76 kW/cm^2 intensity. A 514.5 nm ZnS interference filter with 5 nm bandwidth gave 10 μs switch-on and 20 μs switch-off times with 70 kW/cm^2.

A.Miller, UK

All your work on GaAs has been done on MBE grown material. What's the prospect do you think of using MOCVD grown materials?

Author's Reply

Good, I think. We're going to try an MOCVD sample, so hopefully we'll know before too long. I think the exciton features are there. I think it's a question of whether you can make them flat enough so that you can see nice etalon efects.

Unidentified speaker:

A quest for clarification rather than a question as I'm new to the subject, but why is the power per pixel constant rather than the power density required for the effect. I would expect intuitively for it to be dependent on the power density. Could you perhaps enlarge upon this?

Author's Reply

When one takes an etalon and takes a beam that's smaller than about 25 microns in diameter, as you focus the beam more tightly you find that it takes essentially the same power. You might have guessed that it takes the same intensity, so when one changes that diameter by a factor of three instead of seeing a reduction of a factor of 9 in the required power one finds it's almost exactly the same. What we believe is happening is: when you focus tightly, the carriers (which remove the exciton contribution to the index and give the index change) diffuse out rapidly to a larger diameter. When you try to focus tightly you still fill up the same volume with the carriers. And so in order to supply the same number of carriers per unit time for that volume it takes the same power regardless of how tightly you focus. If you define a small

pixel so that they cannot diffuse out to larger diameters then fine, but it won't go bistable at all unless you focus tightly enough that you put all the power into the pixel because you will still lose them to surface recombination. From that little region's point of view it can't tell the difference whether they diffused out to large diameters or whether they recombined at the surface. But you do get the tradeoff that it switches more rapidly rather than being slow and still requiring high power.

ALL-OPTICAL LOGIC GATES WITH EXTERNAL SWITCHING BY LASER AND INCOHERENT RADIATION

S.D. Smith, F.A.P. Tooley, A.C. Walker, J.G.H. Mathew, M. Taghizadeh and B.S. Wherrett

Department of Physics, Heriot-Watt University, Riccarton, Edinburgh EH14 4AS, U.K.

Summary

Optically bistable, and related, devices have been demonstrated to exhibit input/output characteristics which permit their use as all-optical logic elements and raise the possibility of all-optical signal processing. External switching with both coherent and incoherent radiation has been demonstrated.

1. Introduction

Over the past few years there has been a rapid development in the use of optical techniques for communication and information processing. Most successful of these has been fibre-optics technology, now increasingly penetrating communications networks throughout the world. Other optics based technologies, currently being developed, include:

(i) acousto-optic devices (often in an integrated planar-waveguide format), for radar applications etc.;
(ii) Fourier optics techniques for image enhancement etc., and
(iii) nonlinear convolution or correlation systems, for pattern recognition etc.

These latter techniques all rely upon an analogue processing approach.

A new technology is now emerging based around the phenomenon of optical bistability. This aims at developing all optical digital techniques, equivalent to conventional digital electronics but taking advantage of the ability of a single beam of light to carry independent information in a parallel manner. Furthermore, with the processing rate of microelectronics rapidly becoming restricted by (RC limited) communication times between logic gates, rather than gate switching times, an optical digital processor could achieve unprecedented data handling rates.

2. Optical Bistability

Optically bistable devices rely on some form of optical nonlinearity combined with a mechanism for positive feedback. The nonlinearity can be either absorptive or refractive while the feedback can be provided by an optical cavity, the intrinsic properties of the material (e.g. thermal) or even the nature of the nonlinearity itself. These characteristics combine to give a rapid switch in transmission and/or reflection at a critical input irradiance with, if truly bistable, hysteresis between the switch 'on' and switch 'off' irradiance levels. For this phenomenon to be exploited within the field of information processing, compact solid-state devices operating at low powers are essential. These were successfully demonstrated in 1979 by two groups, working independently, at Heriot-Watt University using InSb (1) and at Bell Laboratories with GaAs (2) (see Smith et al., paper 1, ibid.).

InSb is a particularly useful material in that bistable devices can be made which may be held long-term in either state, thus successfully permitting the demonstration of a range of logic operations based upon the two possible output levels. This contrasts with GaAs in which, as a result of the ratio of nonlinearity to absorption being considerably smaller, damaging thermal effects arise due to the higher power loadings required.

3. All-Optical Logic

An input-(transmitted) output characteristic exhibiting optical bistability is shown in Fig. 1a and, as indicated, it can be seen to act as a memory element. By adjusting the initial detuning from resonance of the device, the curves in Figs. 1b, c and d can be obtained. These single valued characteristics can form the basis of a variety of logic gates. Fig. 1b shows how a single element can act as either an AND or an OR gate, depending on the initial bias condition. (Inverted outputs, giving NAND or NOR logic operations, are obtained from the reflected signal). In addition amplifier and limiter action can be obtained as in Figs. 1c and 1d.

We have demonstrated all these modes of operation using InSb etalons, cooled to 85 K and illuminated by milliwatt power CO laser radiation (5.5 μm) (3). This includes signal amplification in the transphaser operation mode of over 10^4. Recently two bistable elements, spaced 0.5 mm apart on a single InSb chip, have been linked to form a simple optical circuit. Fig. 2 shows the basic layout (lower left). Both gates were operated in reflection and exhibited the input/output characteristics shown (top). The reflection from gate 1 was directed at gate 2, which was already biased close to the switch point. The output from this second gate is plotted against the input to the first (lower right). This overall circuit response, given appropriate biasing of the input gate, demonstrates an XNOR logic function; while the first element is then acting as an XOR gate.

4. External Switching with Coherent Light

In addition to switching these InSb logic devices with 5.5 μm signals a number of other input wavelengths have been employed, while still retaining the CO laser output as the holding beam. In principle any photon energy greater than the bandgap will generate free carriers and initiate switching as a consequence of the resulting refractive index change. We have successfully used 1.3 μm (laser diode), 1.06 μm (Nd:YAG laser) and 1.06 μm (Nd:YAG laser) radiation in this manner (Smith et al. (4)). In the latter example we have shown experimentally (5) that such logic devices can be switched using a 45 picosecond pulse Nd:YAG laser also thus acting as a single pulse detector. The device remains switched into the other state until the holding beam is interrupted. Switching energies of the order of 5 nanojoules have

been recorded and it can be inferred that, since the refractive nonlinearity can be arbitrarily fast according to the rate of excitation, the switch-up speed will be limited only by the build-up time for the optical field in the resonator cavity. With a thickness of the order of 100 μm this is of the order of a few picoseconds.

5. External Switching with Incoherent Light

To take full advantage of the parallel processing potential of an all-optical computer, it is important that direct address with visible images be possible. We have demonstrated that a simple camera-flash directed from outside the cryostat onto a bistable InSb element, will cause it to switch onto resonance.

The external white-light switching pulses were incident on the back face of the crystal. They were generated by a photographic flash unit (Sunpak 3600) and dynamically monitored by a Si-photodiode (Fig. 3). A typical (reflection) characteristic is shown in Fig. 4 for an InSb etalon with L = 260 μm. In the input power range 47 - 51 mW the device had two possible output states. Firing the flash unit caused the resonator to switch from off to on resonance (see Fig. 5).

A critical flash intensity of 39 W/cm^2 was required when the device was biased with 50 mW (1 mW before switch-on point). Assuming a relaxation time of ~ 100 ns for carriers introduced in this manner (5), an effective external energy of 1 nJ is calculated as that required to switch the etalon. The total energy involved in switching the etalon should also include that provided by the CO laser during this time. This total energy is comparable to that observed in other examples of external switching of intrinsic bistable systems by Tarng et al. (6) and in the 1 μm experiments described in the previous section, Seaton et al. (5).

Concurrent with this electronic effect is a thermal one. The absorbed pulse heats the sample. This causes a decrease in the energy gap ($d\tilde{\nu}/dT \approx 2$ cm^{-1}/K at 80 K (7)). Consequently, the refractive index increases (Cardona (8), $dn/dT \approx 6 \times 10^{-4}$) as does both the nonlinear refractive index and the linear absorption coefficient (these changes can be estimated from Miller et al. (9)). The total effect of all three parameters on the input/output characteristic is shown by Fig. 6. Clearly the increase in linear refractive index is dominant in a sample of thickness 260 μm. However, in a thinner sample (L = 98 μm) the nonlinear part of the characteristic is shifted toward the origin suggesting that the increase in non-linear refractive index is dominant.

Shifts of this type can be caused by the heating effect of the flashlamp pulse. For example, Fig. 7 shows that the absorption of 3 mJ/cm^2 (sufficient to raise the temperature ~ 2°K) causes switching of the 260 μm thick device from on to off resonance. That this is a purely thermal effect is concluded from the dependence of the switching point upon only the energy absorbed rather than flash intensity. The 98 μm thick cavity was also switched using this effect. However, in this case the thermal effect causes the etalon to switch from off to on resonance.

6. Visible Light Bistability

An alternative approach is to operate directly with visible radiation. We have recently demonstrated that a multi-layer thin-film interference filter, based on a ZnSe spacer, can give a bistable transmission characteristic (at room temperature) when illuminated with a 35 mW argon-ion laser (514 nm). A mean irradiance of ~ 250 W cm^{-2} is required for switching within the 120 μm diameter focal spot and it responds in a few milliseconds. These figures are consistent with the underlying nonlinearity being thermal in nature. Separate switching by a small additional 514 nm wavelength beam has also been demonstrated.

Thin-film interference filters are particularly attractive as optically bistable or nonlinear devices as their production technology is now well established and large area uniform components can be readily fabricated. Thus two dimensional arrays of all-optical logic elements and novel display devices are amongst the possibilities that could be developed in this manner.

7. Conclusions

Nonlinear Fabry-Perot etalons, with bistable or stepped characteristic, have been shown to act as logic gates and hence provide a basis for all-optical digital computing. Our experimental studies have demonstrated coupling of these gates, a range of external switching possibilities and both infrared and visible wavelength operation. These successfully show at the simplest level, the feasibility of this concept. The infrared carrier wavelength and processing system could readily utilise an array of 10^4 elements in 1 cm^2 of InSb which, switching in ~ 100 ns, implies a data-rate of 10^{11} bits/sec. A total input power of the order 1 W would not all be dissipated in the device. In practice other operating wavelengths, and hence materials, will be necessary to ensure full exploitation of these techniques. With the rapidly expanding data communication rates permitted by fibre-optic networks the advantages of using wide-band all-optical processing techniques in the link interfaces could be considerable. Similarly, all-optical processing of visual information could prove extremely efficient in, for example, robotics applications.

8. References

(1) D.A.B. Miller, S.D. Smith and A. Johnston, Appl. Phys. Lett., 35, (1979) 658

(2) H.M. Gibbs, S.L. McCall, T.N.C. Venkatesan, A.C. Gossard, A. Passner and W. Wiegmann, Appl. Phys. Lett., 35, (1979) 451.

(3) A.C. Walker, F.A.P. Tooley, M.E. Prise, J.G.H. Mathew, A.K. Kar, M.R. Taghizadeh and S.D. Smith, Proc. of Optical Bistability Discussion Meeting, Phil. Trans. Roy. Soc. A (1984).

(4) S.D. Smith, F.A.P. Tooley, A.C. Walker, A.K. Kar, J.G.H. Mathew and B.S. Wherrett, Paper I, ibid. (1984).

(5) C.T. Seaton, S.D. Smith, P.A.P. Tooley, M.E. Prise and M.R. Taghizadeh, Appl. Phys. Lett., 42, (1983) 141.

(6) S.S. Tang, K. Tai, J.L. Jewell, H.M. Gibbs, A.C. Gossard, S.L. McCall, A. Passner, T.N.C. Venkatesan and W. Wiegmann, Appl. Phys. Lett., 40, (1982) 205.

(7) J. Camassel and D. Auvergne, Phys. Rev. B12, (1975) 3258.

(8) M. Cardona, Proc. Int. Conf. Phys. Semicond., Prague; Publ. House of the Acad. Sci., Prague (1960) p. 388.

(9) D.A.B. Miller, C.T. Seaton, M.E. Prise and S.D. Smith, Phys. Rev. Lett., 47, (1981) 197.

Acknowledgements

The authors acknowledge the assistance of Dr. H.A. MacKenzie and Mr. J. Reid in these experiments and the contribution of Mr. N. Ross to sample preparation. The research was supported by both the U.K. SERC and the European Community through the EJOB Programme.

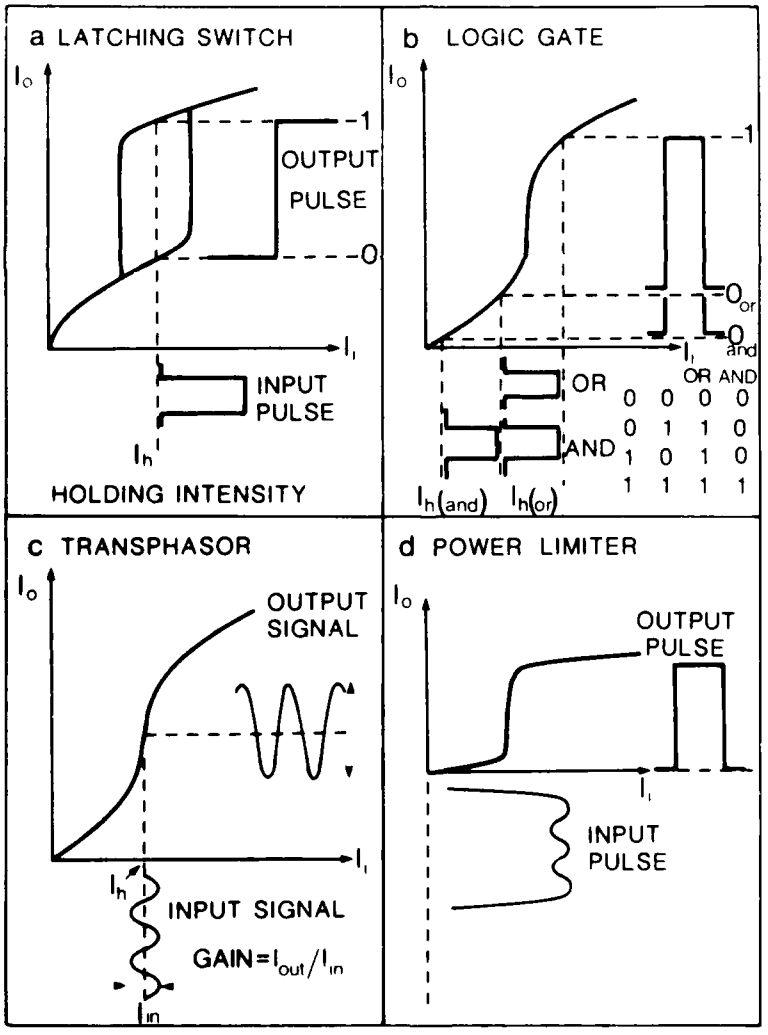


Figure 1. Input/output characteristics for four basic optical circuit elements demonstrated, using InSb etalons.

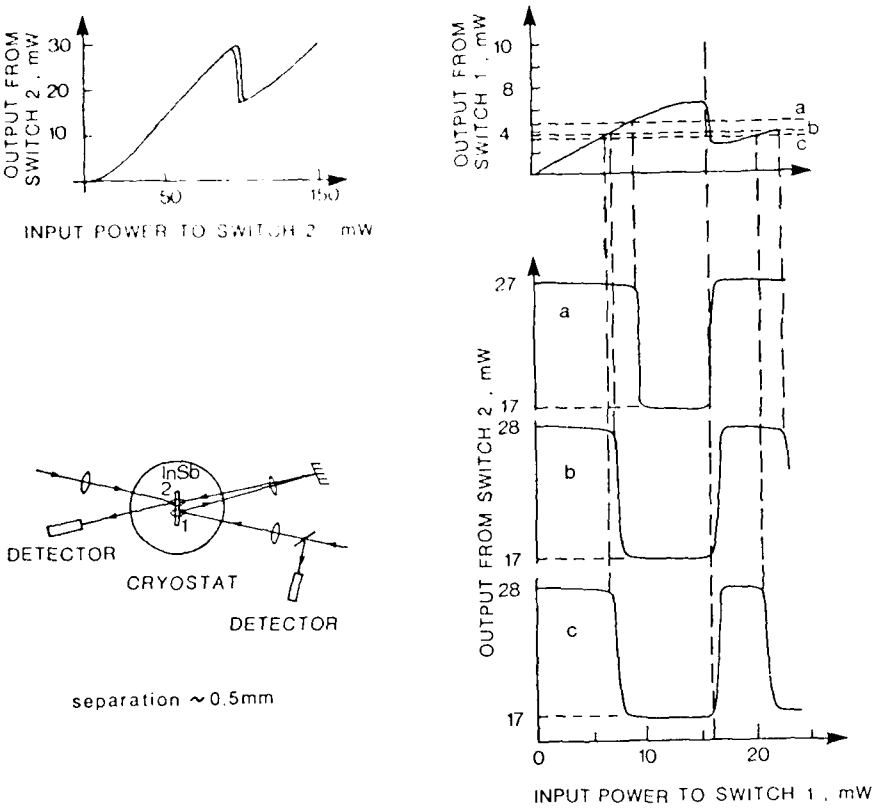


Figure 2. Two coupled OB switches on a single InSb element.
Bottom left: experimental layout.
Top: input/output characteristics of the two switches.
Bottom right: input/output characteristics of the combination.
a, b and c show the result of increasing approach of the initial bias of switch 2 to its switch point.

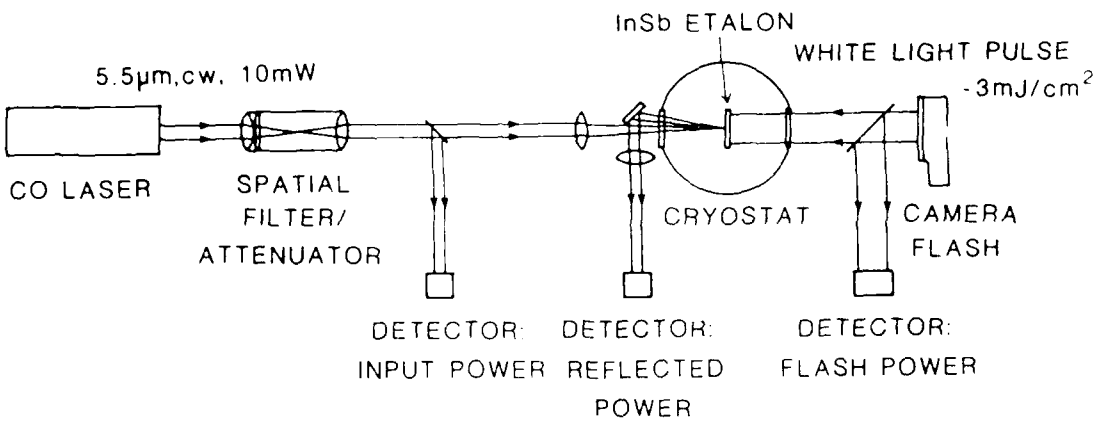


Figure 3. Experimental arrangement used to observe optical bistability and allow introduction of external switching pulses.

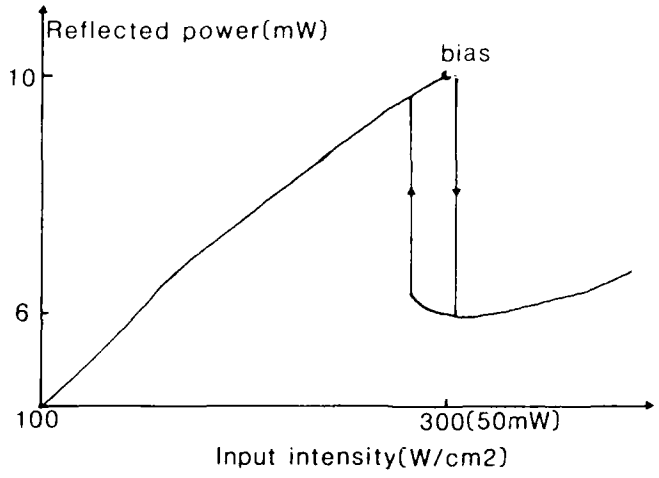


Figure 4. Characteristic of a 260 μm thick etalon showing optical bistability in reflection.

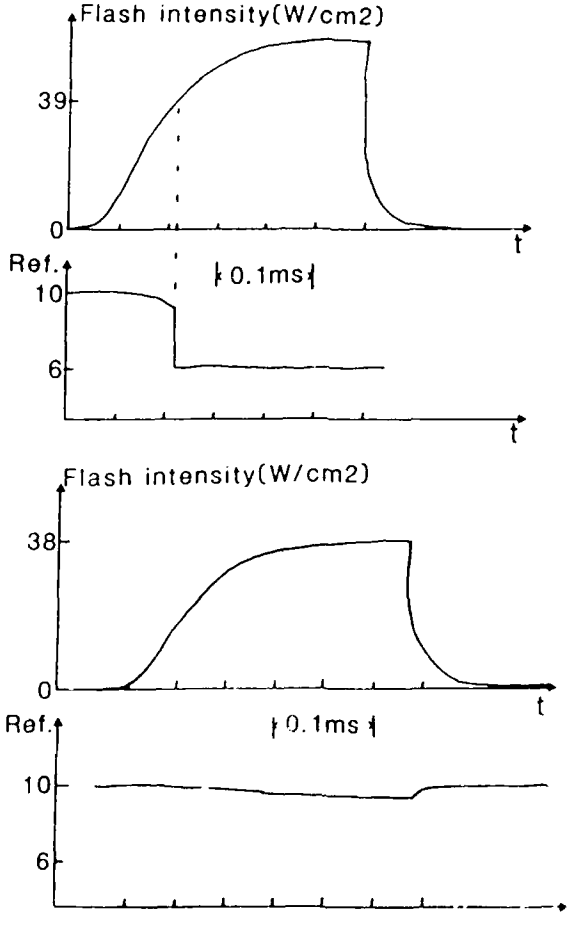
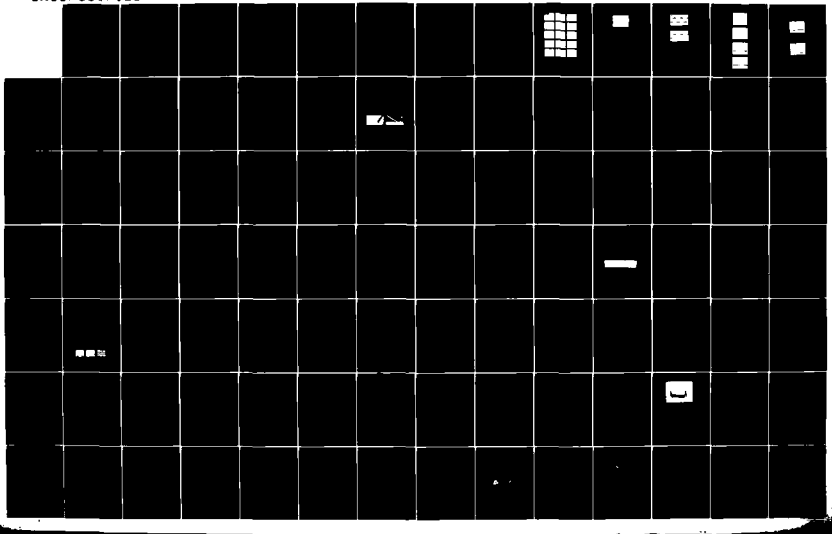
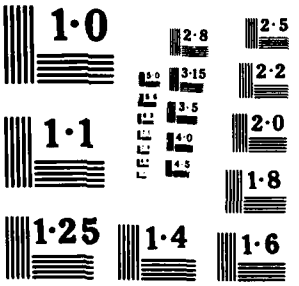


Figure 5. a) and c). The flash intensity incident on the crystal. b) and d). The reflected laser power corresponding to a) and c) respectively. c) and d) show that a high energy pulse will not switch the etalon if it does not reach a critical intensity

AD-A156 908 DIGITAL OPTICAL CIRCUIT TECHNOLOGY(U) ADVISORY GROUP FOR AEROSPACE RESEARCH AND DEVELOPMENT NEUILLY-SUR-SEINE (FRANCE) B L DOVE MAR 85 AGARD-CP-382 1/3

UNCLASSIFIED F/G 20/6 NL




1·0
1·1
1·25
1·4
2·8
3·15
3·5
4·0
4·5
2·5
2·2
2·0
1·8
1·6

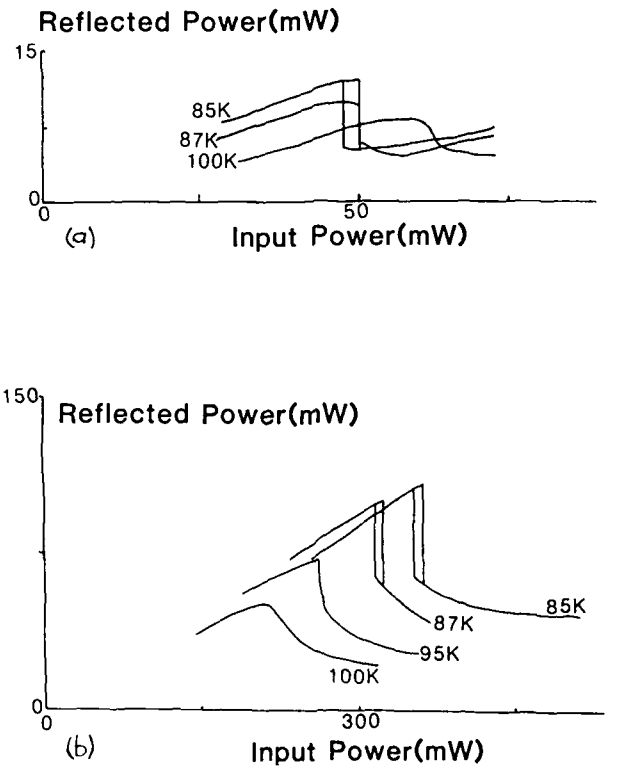


Figure 6. The effect of temperature on the characteristic of a a) 260 μm and b) 98 μm thick etalon.

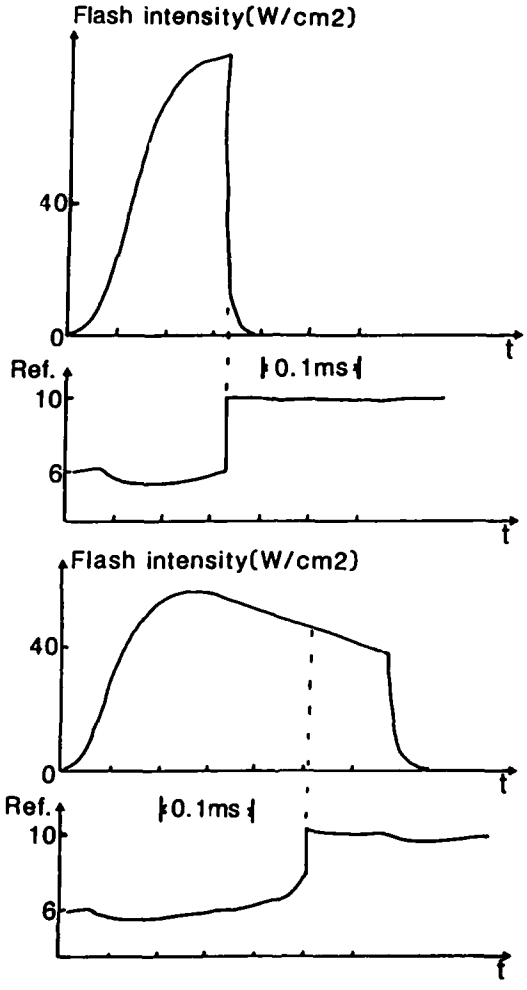


Figure 7. As Figure 5. The etalon is switched from on to off resonance only when sufficient energy is absorbed.

DISCUSSION

J.E.Carroll, UK

Bearing in mind that the nonlinearity is caused by a tensor susceptibility does the off axis or the different axis of address and bias reduce the efficiency?

Author's Reply

The nonlinearities employed in this work are controlled by the third-order nonlinear susceptibility which, unlike the second-order susceptibility, does not require the use of low-symmetry crystals. InSb has a cubic structure, i.e. high symmetry, and our ZnSe thermally evaporated films may have a structure approaching amorphous. Anisotropic effects would not be expected in the experiments we're performing.

H.M.Gibbs, US

What are your switch-on and switch-off times in zinc selenide?

Author's Reply

I missed a viewgraph. Thank you very much. Well there're some simple experiments which we did which just involved ramping the power up slowly and seeing what the switch time looked like. The spot sizes were perhaps large compared with what you might be using. For 100 micron spots, switch-on is about 0.5 ms and switch-off about 1 ms. For 30-μm diameter, they are 150 and 400 μs, respectively. So, as I say, this is the response to just a slow ramp in power up or down in the hundred microsecond region. Clearly if you apply a high peaked power pulse in a short time, it would switch off much sooner.

INTEGRATED ELECTROOPTICAL COMPONENTS USING DIELECTRIC SUBSTRATES

M. PAPUCHON, Y. BOURBIN, S. VATOUX, C. PUECH

THOMSON-CSF Laboratoire Central de Recherches
Domaine de Corbeville, B.P 10
91401 ORSAY FRANCE

ABSTRACT

Since the beginning of integrated optics many electrooptical devices have been demonstrated in a waveguide form. In fact high performances components were obtained using dielectric material as substrates and in particular using ferroelectric $LiNbO_3$ crystals.

After a brief review on the different configurations which have been used to obtain modulators and switches recent progresses will be presented. In addition the use of these electrooptical devices in optoelectronic feed back loops will be described. In particular experiments realized using interferometric modulators have shown the interesting properties of such configurations for optical comparison, linearization of the response of the modulator and differential amplification. Remote control of an electrooptical directional coupler will also be described. Then the use of an optical comparator in a channel of an A/D electrooptical converter will be presented.

I - INTRODUCTION

Integrated optics technology have led to the demonstration of many basic miniature optical components. This is particularly true for the active devices based on the electrooptical effect which gave rise to very high performance devices. As examples it is sufficient to recall that the use of highly electrooptic substrates like $LiNbO_3$, together with the special geometry of integrated optics permitted to obtain very low drive voltage modulators or switches (1, 2, 3, 4) with bandwith practically only limited by electrode design problem.

It is then very attractive to consider these devices as basic elements to realize bistability experiments using optoelectronic external feed back loops. This circuits are to be considered as extrinsic bistable devices compared to intrinsic one where a "direct" optical non linearity is used (Kerr effect for example).

It is well known that, to be able to build a bistable element it is sufficient that a device has an optical output/input transfer function which is non linear with respect to a certain parameter. This remark permits to use pratically all the electrooptical integrated circuit configurations which have been demonstrated to date.

To illustrate this we first describe some of our experiments in bistability using an electrooptic integrated interferometer and a directional coupler switch. Then we will give an example of the use of a bistable device as an optical comparator in an analog to digital converter channel. All the circuits describ here will be realized using Ti indiffusion in $LiNbO_3$

II - INTERFEROMETER WITH FEED BACK

The basic configuration of the device is the standart one and is shown in Fig. 1. All the waveguides are single mode and the electrooptical effect is used to induce phase shifts between the two arms of the Mach Zehnder device. The response of the modulator is a sinusoïdal function of the applied voltage (two wave interferometer) and so is suitable for the observation of bistability effects.

As a basic experiment (Fig. 2) a part of the output light is detected, and the corresponding signal, after an eventual amplification is applied to one of the electrodes of the circuit. A bias voltage is applied to the other electrode to be able to tune the working point of the modulator. The out put of the device is then recorded as a function of the input light intensity for various bias condition. The results are shown on the photograph in Fig. 3. For certain bias voltages the response of the device exhibits the well known hysteresis behavior which is typical of bistable operation. As the modulator response is periodical the same type of response is obtained when the bias voltage increment corresponds to a phase shift of 2π. With such response curve it is clear that many interesting functions can be performed : optical memory, pulse shaping, differential amplification ...

An example of pulse shaping is shown in Fig. 4. The top trace corresponds to the light input signal and the bottom one to the output of the device when it is biaised near the working point corresponding to V = 1V of Fig. 3.

Another interesting feature is the differential amplifier capability. The basic experiment is shown in Fig. 5. The set up comprises the bistable modulator and an optical source (LED in this case) the output of which is sent to the detector. The experimental results are summarized in Fig. 6. When the laser is switched off the signal intensity corresponding to the LED is seen to be 40 mV (Top photograph). When the laser is switched on the small signal from the LED is able to switch on the bistable device leading to an output signal around 2 volts (bottom photograph).

The interferometer is not the only device which can be used for bistability experiments, an example of the use of a directional coupler switch is given in the following section.

III - THE BISTABLE SWITCH

Another device which is usefull in integrated optics is the directional coupler switch. Without entering into the detailed description of the circuit it is sufficient to say that it is a four port device the output of which can be switched via the electrooptical effect through electrical signals applied to control electrodes. The basic experimental configuration is similar to the preceding one (Fig. 7). The light from a laser is coupled in to one of the input part of the switch. Part of the light coming out from one of the ouput port of the device is detected and fed back to the control electrodes (A bias signal can be applied to the switch).

By varying the incident light intensity and by adjusting the bias point of the circuit response curves similar to those shown in the preceding section can be observed. A typical example of these is shown in the photograph of the Fig. 8. The two traces correspond to the two output ports of the switch (vertical axis : output intensity) when the incident light intensity is varied (horizontal axis). Clearly the bistable switching effect is obtained in this case leading to interesting possibilities : for example the output port can be determined by the incident light itself leading to a remotely controlled optical switch (5, 6).

As a last example of our work in extrinsic bistability the use of the bistable device as an optical comparator included in an optical A/D converter channel is described in the following.

IV - A/D CONVERTOR CHANNEL WITH AN OPTICAL COMPARATOR

Integrated optics offers interesting possibilities in optical signal processing. In particular, by using the periodical response of the Mach Zehnder amplitude modulator to an applied voltage, configurations for electrooptical analog to digital conversion have been proposed (7). The main idea is to realize N interferometers on the same substrate with electrode lengths equal to L, 2L, 4L ... The electrical signal to be digitized is fed to all modulators in parallel. It can be shown that, a parallel digital word corresponding to the analog input can be obtained at the outputs of the device if electrical outputs from the detectors pass through comparators which decide if the level is above or below 50 % optical transmission level of the modulators. In this system the laser source is pulsed to sample optically the signal. Holding is not necessary as the conversion time is at the limit given by the transit time of the light in the circuit. Optical comparators can be used instead of the electronic ones this can be achieved by connecting two electrooptic interferometers in a serial way. The second one is used to compare the light intensity coming from the first one and to decide if the level is above or below a predetermined optical threshold.

This experiment has been performed in our laboratory using two Mach Zehnder interferometers connected in series. The light from a semiconductor laser is pulsed at the sampling rate of the A/D channel and the signal to be digitized is applied to the first modulator. The second modulator is inserted in a feed back loop the input light being the output of the first one.

The photographs of the figure 9 show a typical result. The top one corresponds to the output of the laser when pulsed at 100 KHz. The middle one is the output of the first modulator when a symetric sawtooth voltage is applied to it with a peak to peak voltage equal to V_π . We recognize the sinusoïdal response of the interferometer.

The bottom photograph corresponds to the output of the optical comparator when it is biased to a threshold value of 50 % of the maximum input signal.

Photographs of Fig. 10 show the same kind of results with sampling frequencies of 150 kHz (top photograph) and 1 MHz (Bottom photograph).

V - CONCLUSION

We have given here some examples fo extrinsic bistability obtained on our laboratory using integrated optic electrooptic devices. First we started by using a Mach Zehnder Interferometer/modulator in a feed back loop. Various kind of functions can be obtained : optical memory effect, pulse shaping, differential amplification ... Then the directional coupler is shown to be also suitable for these experiemnts leading to interesting possibilities like remotely controlled switches. At last the bistable interferometer is used as an optical comparator in one channel of an electrooptical analog to digital converter.

The authors thank J.M. ARNOUX, A. ENARD, D. PAPILLON and M. WERNER for technical assistance.

These studies have been partly supported by DRET and ESA.

REFERENCES

1 Y. BOURBIN, M. PAPUCHON, S. VATOUX, J.M. ARNOUX and M. WERNER, "Polarization independent modulators with Ti : $LiNbO_3$ strip waveguides", Electron. Letters, 20, June 1984, p. 496-497.

2 R.C. ALFERNESS, S.K. KOROTKY, L.L. BULH and M.D. DIVINO, "High speed low loss low-drive power travelling wave optical modulator for λ = 1,32 μm", Electron. Letters, 20, April 1984, p. 354-355.

3 C. DUCHET and R. MARTIN, "Electrooptic modulator on Ti:$LiNbO_3$ with very low drive voltage", Electron. Letters, 20, June 1984, p. 567-568.

4 P.S. CROSS, R.A. BAUMGARTNER and B.H. KOLNER, "Microwave integrated optical modulator", Appl. Phys. Lett., 44, March 1984, p. 486-488.

5 P.S. CROSS, R.V. SCHMIDT, R.L. THORNTON and P.W. SWITH, "Optically controlled two channel integrated optical switch", IEEE J. Quantum Electron., Vol QE-14, August 1978, p.577-580.

6 A. SCHNAPPER, M. PAPUCHON and C. PUECH, "Remotely controlled integrated directional coupler switch", IEEE J. Quantum Electron., Vol QE-17, March 1981, p. 332-335.

7 H.F. TAYLOR, M.J. TAYLOR and P.W. BAUER, "Electrooptic analog to digital conversion using channel waveguides modulators", paper Tu C1 Technical Digest of Topical Meeting on Integratel and Guided Wave Optics, Salt Lake City (1978).

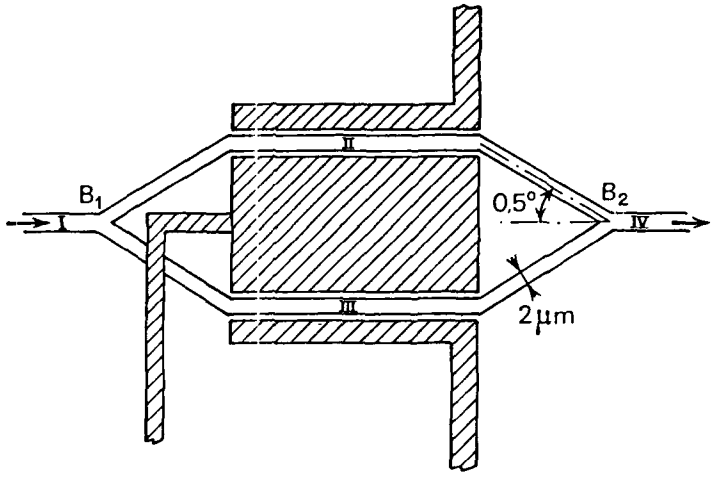


FIGURE 1

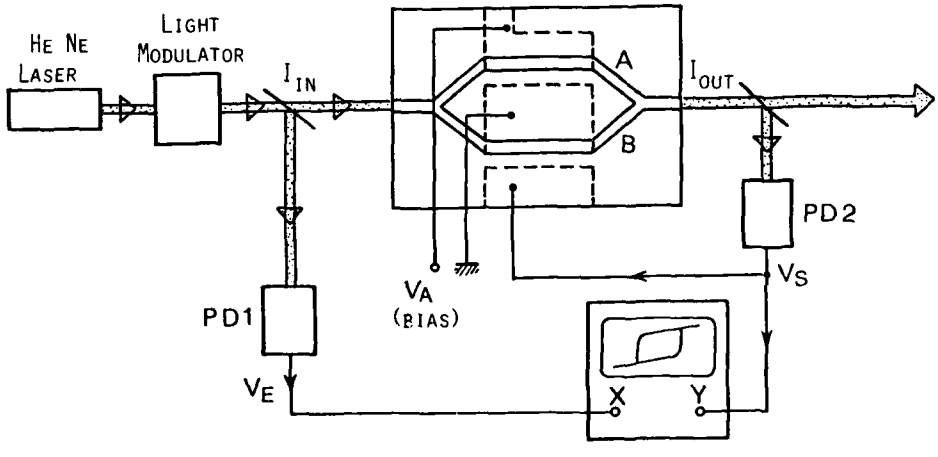


FIGURE 2

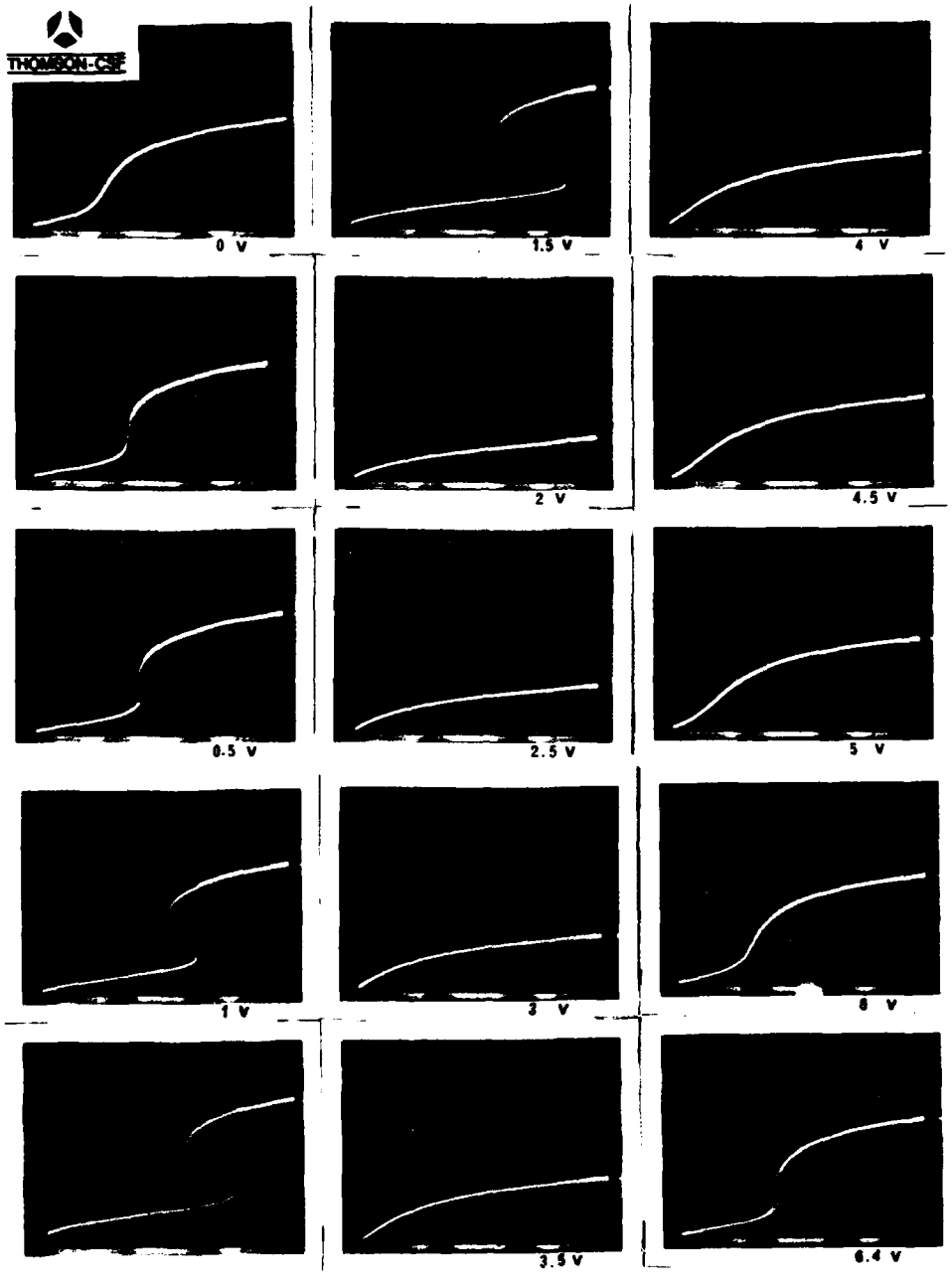


FIGURE 3

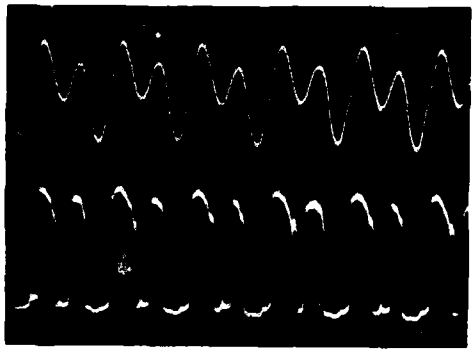


FIGURE 4

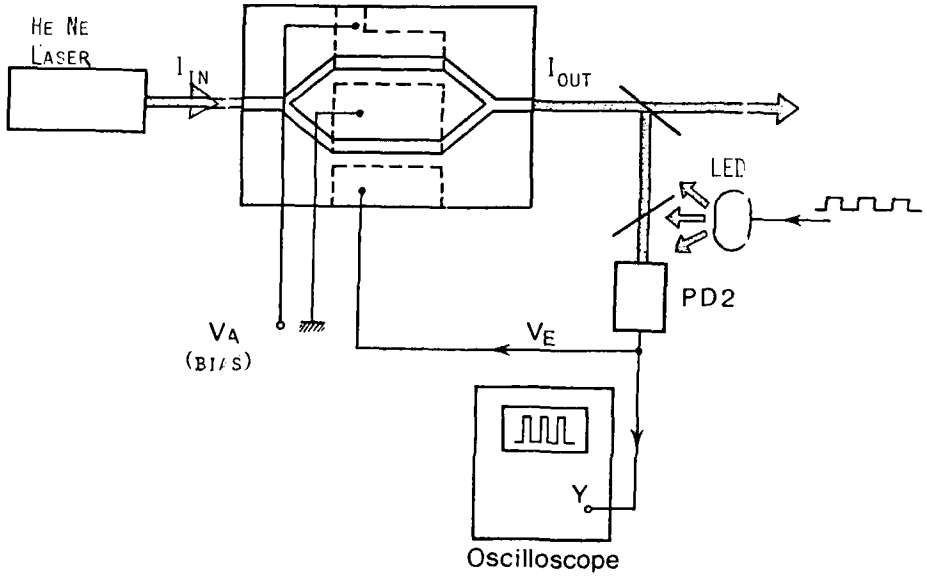


FIGURE 5

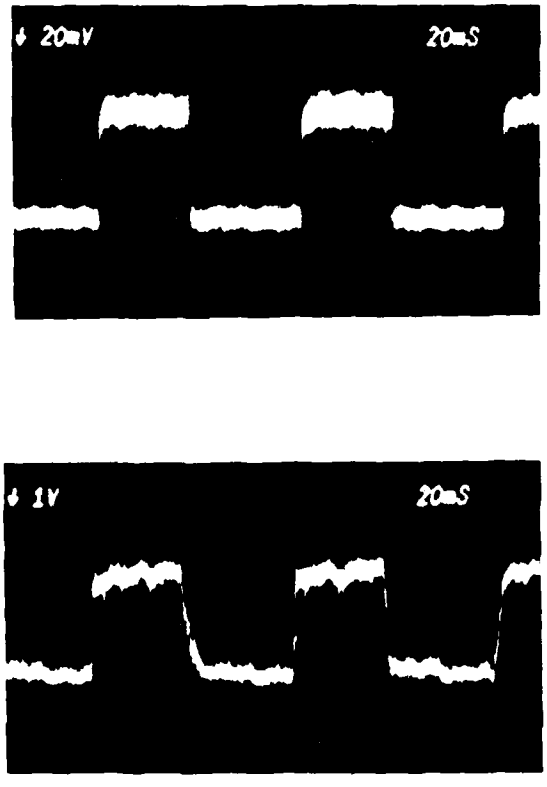


FIGURE 6

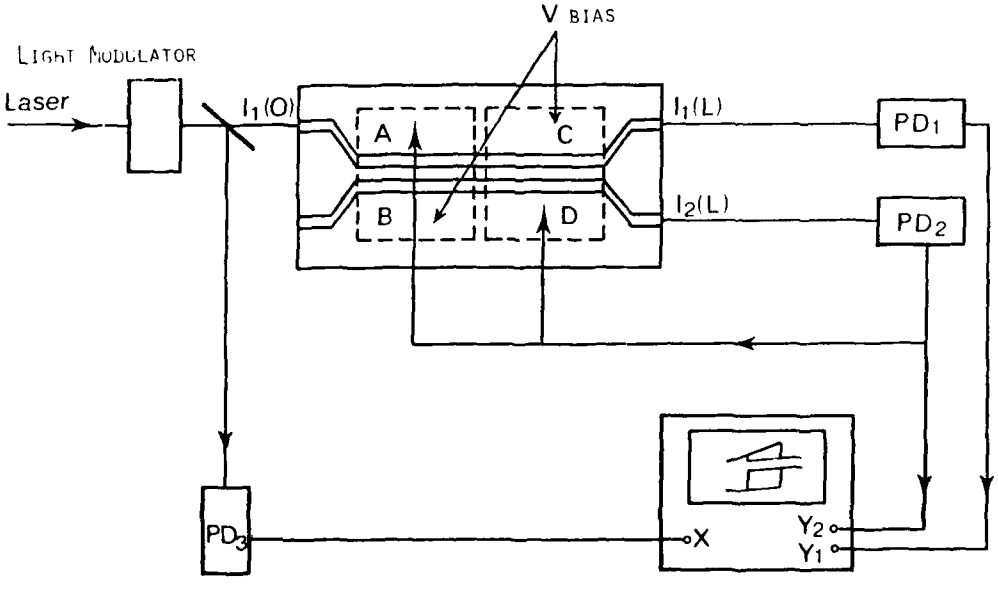
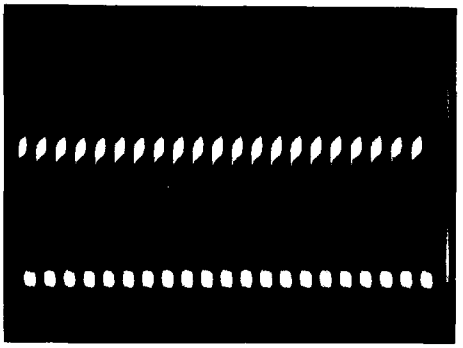


FIGURE 7

FIGURE 8



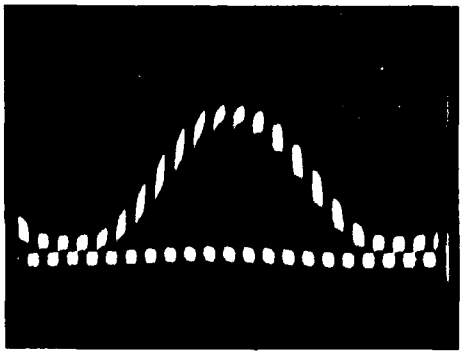
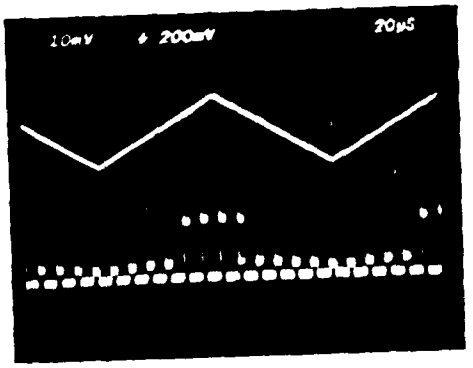


FIGURE 9



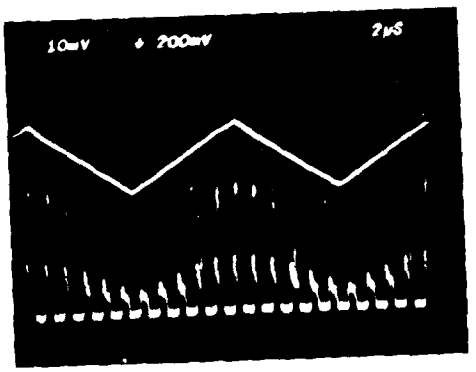


FIGURE 10

variations for different devices (eg the curved output guides of a rib waveguide directional coupler may require extra etching to increase confinement (17)).

The inclusion of a laser on the chip considerably increases its usefulness. Discrete semiconductor lasers have cleaved facets acting as mirrors, and this is not really compatible with integration. However, there are other methods for making semiconductor lasers which do not involve conventional cleaving, and some of these are more suited to integration.

(i) Etched facets. High quality facets have been obtained using etching techniques (86), but there will always be problems in coupling the output from an etched facet laser to a monolithically integrated optical waveguide. Fresnel reflections are the main difficulty, - reflection loss will limit the coupling efficiency, and the reflected light will cause spectral fluctuations. However, when a laser needs to be coupled to an external guide (fibre), or when a low coupling efficiency can be tolerated in an integrated circuit (eg coupling to a monitor photodetector) etched facets may be useful.

(ii) Micro-cleaved facets. In this technique, the active layer is exposed as a cantilever by selective etching, and then removed ultrasonically. Very low threshold current lasers have been obtained by this method (87). They have the same features as etched facet lasers for monolithic coupling to waveguides.

(iii) Ring laser (88). In principle this is a highly integrable laser, but it requires an output coupler. The problem of achieving a ring structure with sufficiently small radius of curvature (<100μm) such that threshold currents are low together with achieving low radiation loss from the curved waveguide is a major limitation to this method

(iv) Distributed Feedback (DFB) laser (89). In this device a grating is incorporated in or near the active region of the laser; this acts as a distributed reflector and mirrors are not required. Although the technology for making these complex devices (the grating period needs to be of the order of 0.25μm) is becoming reasonably well established, a difficulty arises if one wants to integrate the laser with a waveguide. This is because the laser active region, although transparent when pumped, is absorbing outside the gain region. Thus a more complex structure is needed. Fig 4 shows a simple approach in which the basic three layer DFB structure is grown, prior to etching and subsequent growth of a p InP layer over the whole structure. The lower InGaAsP laser confining layer becomes the waveguide layer, and proton isolation could be used to isolate electrically the waveguide from the laser. In this simple structure, the overlap between the optical fields is not optimum, but improved structures may be envisaged - although they will require even more complex epitaxial growth.

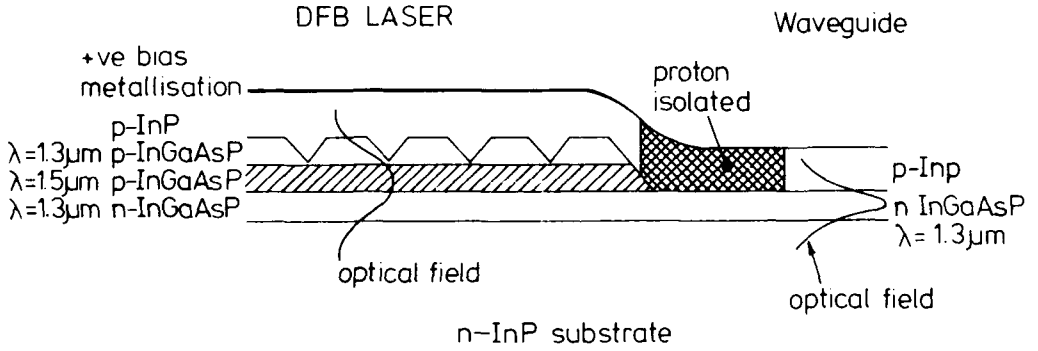


FIG 4 A scheme for integrating a DFB laser with a waveguide

(v) Distributed Bragg Reflector (DBR) (90). The structure can be similar the the DFB, but the gratings are positioned outside the gain region in the waveguide layers.

It is the latter two techniques that have received most attention - devices with gratings incorporated have advantages as discrete devices because they operate with a single longitudinal mode. The DFB is probably the easiest to make - and can also be considerably shorter.

Aiki et al (91) integrated 6 GaAlAs/GaAs DFB lasers of different wavelengths onto a single chip, using waveguides to combine the signals into one output waveguide. The overall efficiency was poor, but this is the most ambitious demonstration of the integrated optics concept. Mertz et al (92) have coupled etched-facet lasers to detectors via straight and curved waveguides, and although losses were once again rather high, these circuits do demonstrate the feasibility of such integration, and with improved designs and fabrication techniques, better performance should be possible.

Although there has been limited work on the integration of lasers and waveguides, there has been very significant progress in integrating electronic devices with lasers. In one

applications, but could be important if the incident optical power is limited, such as in telecommunications applications. In principle, however, it should be possible to combine a short lifetime with a reasonable mobility, and this must be a technological challenge for the future. Fast photoconductive detectors have been made in GaAs (58), and InP (69), with response times typically less than 50ps.

This approach is particularly attractive, because a photoconductor is an extremely simple structure (merely an optically-sensitive resistor) and it should be possible to integrate these with great ease. However, other detectors may be suitable. In particular, a GaAs Schottky device with a response time of 5.4ps and a bandwidth of 100GHz has been demonstrated (70) and InGaAs/InP PIN diodes have shown response times of 30ps (71). These devices are more complex than the photoconductor, and require epitaxial layers. However, they may give advantages in responsivity and sensitivity.

7 NOVEL STRUCTURES

The development of MBE (molecular beam epitaxy) and MO-VPE (metallorganic vapour phase epitaxy) which can controllably produce very uniform multilayer structures of thicknesses from a few atomic layers to several microns, has allowed the fabrication of several novel structures which are likely to facilitate the development of optical circuits.

The quantum well (QW) structure (72) consists of a narrow bandgap layer sandwiched between wider bandgap layers, the narrow layer having a thickness less than the de Broglie length (≈100Å), so that carriers are confined to two-dimensional motion. This results in the density of allowed states becoming a staircase function, bounded by the parabolic curve for bulk material. This, and other differences, influence a large number of the fundamental properties of the material such as the gain and absorption spectra, transport properties and various non-linear effects. Often, many quantum wells are stacked together (multi quantum wells, MQWs) to enhance these effects.

A variant on the MQW is the strained layer superlattice (SLS) (73). In conventional epitaxy, all the layers are usually grown lattice-matched to the substrate to prevent strain and dislocations. In strained layer superlattices the restriction that all the layers must be lattice-matched is relaxed, but the layer thicknesses and mismatches are controlled so that the overall strain is minimised or nulled such that the individual strains are kept below the critical value at which defect formation would occur. SLSs are often, but not always, MQWs, and thus have all their benefits, but additionally offer a greater degree of flexibility on the choice of compositions that may be grown on a given substrate. Perhaps more significantly, they add an extra degree of freedom which could be used to optimise the properties of the effective band structure of the composite material. Details of the conduction band and valence band edge might be changed, and also effective masses and thus transport properties. Additionally, the band structure can become sufficiently distorted by the local strain, that direct bandgap superlattices can be made using indirect bandgap constituents (eg GaAsP on GaP) (74).

QW structures in GaAlAs/GaAs have already been used to demonstrate lasers with reduced threshold currents (75), reduced and thickness controlled lasing wavelengths (77), sharper gain spectra (76), and are expected to show other benefits (78-80). However, because of the effects of non-radiative recombination, particularly Auger recombination (81), it is not clear yet how many of these benefits will transfer to other materials systems, particularly those for longer wavelength applications, where Auger recombination is more important. The enhanced gain in MQW lasers will allow reduced current operation, thus lower thermal dissipation, and allowing the possibility of integrating a large number of lasers on one chip. Threshold currents of only a few mA have already been demonstrated, and further reductions are possible. Another major benefit of MQW structures is enhanced non-linear effects as discussed in section 5.

Superlattices in which there are periodic doping, rather than compositional variations, (the nipi superlattice (82)), offer a new range of tunable properties including optical gain (83). The tunability, and the fact that the gain operates below the band edge could result in a simpler technology for integrating lasers and waveguides, as well as matching emission wavelengths to particular resonances for non-linear devices, or in matching local oscillator and signal oscillator frequencies for coherent optical systems (84).

Thin epilayers have also allowed the development of new electronic devices, for example the use of a two-dimensional electron gas in the high electron mobility transistor (HEMT) (85). This should benefit integrated optics because of its improved integrability and performance.

8 INTEGRATION

Semiconductor optical integrated circuits offer the possibility not only of a wide range of optical processing functions, but also opto-electronic and electronic devices on the same chip. The technology to realise these concepts appears formidable at present (see section 9), but progress is rapid, and a start at demonstrating limited degrees of integration has been made.

To date, the passive and electro-optic features have only been developed to the 'building block' phase. More advanced components such as star couplers, matrix switches, A-D converters and others as demonstrated in $LiNbO_3$ have not been made in semiconductors yet, although with recent improvements in epitaxial techniques and the fabrication of relatively low-loss, low voltage waveguide devices recently, progress could be rapid. In many cases a single epitaxial layer structure could be used, perhaps a variant on fig 2, but with slight processing

This pulse actually consists of many external cavity modes locked in phase, and the Fourier transform of such a composite waveform is a train of pulses of width approximately t_r/N and period t_r - N is the number of modes locked together. With a typical gain width of 100nm at 1.55μm, pulses as short as 30fs could, in principle, be generated if mode locking occurs over the whole spectrum. In practice, unwanted sub-cavity modes and, ultimately, dispersion prevent this. Bandwidth limited ($\Delta\tau x\Delta\nu$=0.36) pulses of 16ps have been obtained (58). The external cavity needs to be quite long (several cm) to allow conveniently low pulse frequencies, and in one case (65) a SELFOC resonator has been used as the cavity. Active modelocking using semiconductor lasers has also been demonstrated by driving the laser cw and using an optical switch within the external cavity. Using an InGaAsP/InP laser with a $LiNbO_3$ travelling-wave modulator, pulses of 22ps have been generated at 7.2GHz (36).

In passive modelocking, a laser with a region of saturable absorption, is placed in an external cavity and driven with a dc bias only. Although the mechanism is not completely understood yet, the evolution of picosecond pulses can be considered as follows (59); if the light intensity exceeds a critical value it can saturate the absorber so that the most intense light sees gain on both passes through the semiconductor. The gain will become depleted, thus truncating the light pulse, but the absorber bleaches more rapidly than the gain depletes providing a short period of net gain. This process will continue during each round trip with the pulse gaining in amplitude but shrinking in width until it is limited by dispersion. Van der Ziel (59) obtained 0.65ps pulses, which was consistent with dispersion limitations. These broadened to 20ps bandwidth limited pulses when an etalon was introduced. The exact properties of the saturable absorber are rather important, and most of the work seems to have been carried out with degraded or damaged semiconductor material. However, by using a semiconductor laser with a split electrode configuration (64) the amount of saturable absorption may be controlled reasonably well.

There has been a large number of publications describing lasers which unintentionally produce sub-nanosecond pulse trains at rates between 0.1 and 1GHz when driven with a dc current. There are probably several different ways in which this behaviour could arise, eg saturable absorbers, deep traps or non-uniform excitation in the active layer. As, at present, these effects are uncontrolled, they will not be considered as useful sources of picosecond pulses.

Picosecond pulses can also be obtained by optically modulating the output of a cw laser. Using a comb-generator driven interferometric travelling-wave $LiNbO_3$ modulator, pulses of 45ps have been achieved (60), and 47ps has been demonstrated using a $LiNbO_3$ travelling-wave directional coupler (66). Using a standing wave resonator to avoid the limitations set by the differing microwave and optical propagation velocities in $LiNbO_3$, pulses of 19ps at 20GHz have been produced (67). However, being a resonant structure this type of device will only operate within a small range of pulse frequencies near the resonance. As mentioned earlier, semiconductor travelling wave devices may allow higher repetition rates and shorter pulses. Short pulses can also be generated by overdriving optical modulators (61,62) but all these methods have the disadvantage of dissipating the unwanted energy. Thus the output energy per pulse is likely to be rather small.

It is clear that there are several methods of obtaining pulses from 0.5 to 50ps from semiconductor devices at frequencies of 10^5 to 10^{10}Hz. Some of these methods, in particular modelocking, Q switching and gain switching, are capable of high peak powers which will be useful for non-linear applications. Although these methods are useful for providing discrete optical pulse generators, their usefulness, particularly for optical logic, picosecond sampling and optical communications will be considerably enhanced with monolithic integration. In principle, lasers may be integrated (see section 8) either by using on-chip facet techniques, or by using gratings, so it should be possible to realise these pulse generators in integrated form. The integration of techniques 1,3,5 and 6 in table 2 is conceptually easy, but for modelocking a relatively long external cavity is required, together with some form of bandwidth restriction. One approach would be to use a DFB laser, which would act as the source and a method of controlling the bandwidth, with an integrated low loss waveguide as the cavity. A long cavity will be needed if relatively low pulse repetition rates are required (5cm in a semiconductor corresponds to a repetition rate of about 1GHz). Long waveguides may be realised using curved sections to increase the packing density. Clearly there will be considerably less problems at the higher bit rates where cavity lengths may be shorter. Dispersion could be the dominant limitation to short pulses in such structures.

6.2 Pulse detection

As well as optical pulse generation, it may be necessary to convert fast optical pulses to electrical pulses in a semiconductor integrated optical circuit. Ideally, of course, it would be desirable to avoid completely the need for high speed electrical pulses - ie the various optical processing stages should ultimately provide an electrical output at a low speed which could be handled relatively easily. Indeed, the transmission of picosecond electrical pulses without crosstalk and other forms of corruption is a major problem in its own right, and one of the reasons why optics is so attractive for high speed applications. It would appear, however, that elements for detecting picosecond optical pulses will not be as difficult to make as those to provide the pulses in the first place, and their integration should be relatively easy.

The most studied detector for ultra-fast pulses is the photoconductor. The speed of these devices is determined by the carrier lifetime, which is influenced by recombination in the bulk or at interfaces, surfaces or contacts. Photoconductors can show gain, and thus high responsivities, but the methods of reducing lifetime often employed, (eg proton damage), usually reduce the mobility and thus the gain. This might not be a problem for many

	pulse width (ps)	limit (ps)	pulse rate (Hz)	comments
1 Gain switching - pulsed - sinusoidal	5 - 30 30	τ_{ph} (2L/c)	single →2×10^9 10^8 → 10^{10}	simple, on demand pulses, multimode or single mode
2 Optically pumped gain switching	1	τ_{ph}	≈ 10^9	very short cavity (few μm) usually with high reflectivity, tunable using thickness variation
3 Q switching (ie loss switching)	≈14	τ_{ph}	3×10^9→ 1.2×10^{10}	loss and gain regions may be integrated, on demand, high power pulses
4 Modelocking -active -passive	5.3 .56	$\frac{1}{\text{(gainwidth)}}$ sub psec		periodic, needs high quality ext cavity (very good ar coatings), integrated external cavity demonstrated
5 Self pulsation	20 - 100		10^8 - 10^9	poorly understood and uncontrollable
6 Optically switched cw optical input	19		10^9 - 10^{11}	conceptually very easy to integrate, low peak power

TABLE 2 Different methods of obtaining picosecond pulses from semiconductors

demonstrated with both GaAlAs and InGaAsP lasers, using avalanche transistors, step recovery diodes and rf drive. Pulses as short as 15ps have been obtained, and shorter pulses predicted (53). For methods where the electrical pulses are available on demand, the optical pulses will also be available on demand - although some patterning will arise for pulses closer than the recombination time (a few ns). The output pulse of such lasers is spectrally impure, having many longitudinal modes, but this can be overcome by using external cavity mode selection methods such as a grating (54) or possibly by switching a DFB laser. The ultimate aim would be for Fourier transform limited linewidth / pulsewidth products. One of the advantages of such gain switching is the very high peak powers that can be obtained (80 - 20,000mW), although it is unlikely that mean powers would exceed typical cw laser power capabilities.

The gain switching may also be obtained by optical excitation, and this method has been investigated extensively with platelet or film lasers in which a thin layer of semiconductor (2-20μm thick) is placed between two reflectors and excited by picosecond optical pulses from a mode-locked non-semiconductor laser. The gain width can be very broad (several kT), but with only one Fabry-Perot mode falling within this gain width, single longitudinal mode emission is obtained. By using wedge shaped samples, the wavelength of the mode can be tuned over the whole gain width (55), and by using the appropriate bandgap semiconductor, a very wide range of wavelengths may be produced.

In the Q switched method (56), a gain and switchable loss mechanism are incorporated within the cavity. The gain region is driven hard to a large value of gain, but lasing is prevented by the lossy region. If this loss is now decreased, a lasing pulse will build up very rapidly. This pulse will terminate itself by the rapid depletion of the gain. The loss region should return to a high loss state before the gain can build up. Both the gain and loss regions have been integrated on the same chip (63) by using split electrode configurations. Pulses of <40ps (detector limited) at a repetition rate of 14GHz were obtained using an InGaAsP/InP structure. Peak powers of several hundred milliwatts should be obtained using this method.

In active mode locking (57) a semiconductor laser with one perfectly anti-reflection coated facet is placed in an external cavity. The laser is modulated with a sinusoidal rf or pulse train having a modulation period equal to the photon round trip time (t_r). The emitted pulses see gain on arrival at the laser, and the peak sees the highest gain thus sharpening the pulse.

5 NON-LINEAR AND BISTABLE ELEMENTS

A range of devices employing non-linear processes may be envisaged. In particular, in the same way that the whole of electronics is based on non- linear electronic devices, a similar family of optical devices may eventually have applications.

Non-linear optical effects are caused by the non-linear polarisation of the medium through which the optical wave is propagating. Various non-linear effects corresponding to the different order susceptibility terms have been investigated in a variety of materials. Although the second order susceptibility term (responsible for second harmonic generation, parametric oscillation and amplification) can be quite large for semiconductors compared with other materials (40), most of the work on III-V semiconductors has concentrated on the third order susceptibility ($\chi^{(3)}$). This gives rise to an intensity dependent refractive index which is particularly relevant to optical logic applications. Optical bistability (OB) may be obtained if a material with an intensity dependent refractive index is placed in a suitably designed Fabry-Perot cavity (41). The value of $\chi^{(3)}$ in III-V semiconductors using wavelengths well away from the band edge (the non-resonant case) is extremely small, and will not be particularly useful because extremely high powers will be needed, even with the small active volumes offered by waveguide structures. Most of the work on OB has involved working near an absorption resonance. Under these circumstances $\chi^{(3)}$ can be enhanced quite dramatically. OB associated with resonant effects has been observed in GaAs (42), GaAs/GaAlAs multi quantum well structures (43), InAs (44), and InSb (45). In the latter material, a range of optical logic elements have been demonstrated including AND and OR gates, and a "transphasor", the optical analogue of a transistor. The power density for operation of some of these InSb devices has been extremely low ($10W/cm^2$ at a temperature $\leq 77K$). However, InSb is probably not the optimum semiconductor material for OB devices; there are no ternary or quaternary compounds which can be grown epitaxially on InSb for an optimum waveguide structure and/or the integration of other optical and electronic components. In addition, the band-gap is probably too narrow for use at useful wavelengths at room temperature. However, the work on InSb has demonstrated that devices employing OB are certainly feasible.

Whereas the non-linearities in InSb are due to effects associated with transitions between conduction and valence band edges, those in GaAs are thought to be associated with exciton states. The preliminary work on simple GaAlAs/GaAs/GaAlAs etalons showed bistability at low temperatures, but not near room temperature where the exciton binding energy becomes comparable with kT. However, OB has been demonstrated at room temperature in GaAs/GaAlAs superlattices - the exciton binding energy being increased in such structures. Although power densities are quite high ($10^5W/cm^2$), scaling to waveguide dimensions indicates that power levels will be well within those available from semiconductor lasers. In fact semiconductors with appropriate focussing have been used to demonstrate OB at room temperature in the superlattice structures (46).

A major problem with the operation of these devices at present is their speed. The $\chi^{(3)}$-induced OB depends on carrier dynamics, and the turn off-time depends on the carrier recombination time (a few nsec in normal material at room temperature). Although this could be reduced by modifying the material appropriately, eg introducing fast recombination centres, the switching power would then be compromised. Indeed, it seems there will be a trade off between switching power and speed, and more work is needed to establish how favourable this is for device operation - especially waveguide devices.

An alternative approach to obtaining bistable effects is to use a hybrid arrangement (51). The optical output from a Fabry-Perot cavity is detected, and used to generate a voltage which varies the refractive index within the cavity by the electro-optic effect. This sort of feedback system allows a non-linear light input / light output relationship, hysteresis and OB. The electronics can have gain, so very little power is needed to demonstrate OB, and the cavity medium need not have an intensity dependent refractive index. The speed of such devices is limited by that of the electronics. As one of the major benefits of optical logic is its potential for ultra-high speed, this hybrid approach may not be the most attractive option in the long term. The ability to integrate all the optical and electronic elements on one chip could, however, be very useful.

6 PICOSECOND PULSE TECHNIQUES

6.1 Pulse generation

There has been considerable interest in generating very short optical pulses, either as a repetitive sequence, or 'on demand' (ie with variable pulse separation), for a wide variety of applications including fast optical signal processing, picosecond spectroscopy, sampling and optical communications. Semiconductor lasers have been used in a variety of configurations to produce such pulses (see table 2) and can have the merits of compactness, easy operation, high repetition rates (many GHz), tunability and a wide range of wavelengths. Most of the the work to date has focussed on demonstrating and understanding the techniques. This section will describe the basic methods of producing picosecond pulses using semiconductors, and comment on some aspects of integrating picosecond pulses with other components.

The main methods of obtaining short pulses using semiconductors are summarised in table 2. The simplest of these is to drive the laser with a short but intense current pulse which rapidly raises the carrier concentration from well below threshold to well above it. The output power builds up rapidly until it starts to deplete the carriers and thus reduce the gain. If the drive pulse is terminated at this stage, a short high power optical pulse is obtained; if the drive pulse is continued, the damped relaxation oscillation transient characteristics typical of injection lasers will be obtained. This method of 'gain switching' a laser has been

Device	Length (mm)	Voltage (V)	Bandwidth (GHz)	Loss	λ (μm)	$\frac{V_\pi \lambda}{\lambda^2}$	ext ratio (dB)	ref
GaAs/GaAlAs								
PM	4.2	9	>1.2 (est)	5.5dB/cm	1.15	29		32
DC	9	15		1dB/cm	1.56	55	15	9
DC	3.5	40	3	8-10dB (L-F)	1.3	83	15	6
MZ	2.5	25	5 (est)		1.3	37	14.5	33
MC	4.9	12.5			1.06	55	27	13
InP								
DC	6.5	12			1.51	34	16	25
$LiNbO_3$								
DC		16	2	2dB (F-F)	1.5	64	18	34
DC	9	11	1.5	4dB (F-F)	1.5	44	21	34
DC	16	6		2dB (F-F)	1.3	57	27/21	35
DC(TW)	14.5	4.5	7.5	8dB (F-F)	1.3	39		36
DC(TW)	9.5	7.5	10.5	7dB (F-F)	1.3	42		36
DC(TW)	19.5	4.5	5.5	2.2dB (F-F)	1.3	52		36
DC(TW)	9.5	18	10	1.5dB (F-F)	1.3	101		36
MZ	4	4	2.75		0.85	22		37
MZ(TW)		4	13.2		0.85			37
MZ	3	4.5			0.63	33	17	38
MZ	6	3.8	1.1		0.65	54		39
MZ(TW)	4	2	13		0.84	11		47
MC	5	17			1.35	47	17	11

Notes

1) PM=phase modulator; DC=directional coupler; MZ=Mach-Zehnder;
MC=TE-TM mode converter;
TW=travelling wave electrode configuration

2) F-F corresponds to a fibre-to-fibre insertion loss measurement
L-F corresponds to a semiconductor laser-to-fibre insertion loss measurement

TABLE 1 Comparison of various electro-optic waveguide devices made with different materials

Photoelastically induced waveguiding in semiconductors has also been studied (49,50). In these structures, the refractive index is varied by introducing strain by depositing a metal or dielectric layer on the semiconductor. These waveguides have the advantage of ease of fabrication, and being planar, such guides may be affected less by scattering. However, unlike other techniques, lateral and vertical confinement cannot be controlled independently, and there are uncertainties about the long term behaviour of devices which rely on stress effects.

Perhaps a more promising structure for the future is the buried waveguide (27,28). In this type of structure, the guiding region can be grown in a groove etched in the substrate or epilayers (27) - a technique commonly used for buried heterostructure lasers. Alternatively, the guides may be buried by growing over an etched rib waveguide (28). These techniques, usually requiring two stages of epitaxial growth, are more complex than the simple rib waveguide, and so far the losses have been >5dB/cm. However, buried guides are expected to give lower scattering losses than those encountered in exposed rib structures, and they are much more compatible with many of the advanced laser structures. Thus they may be more suitable for integration.

For many applications InP-based materials will be preferred, - particularly for structures in which lasers and detectors operating in the wavelength range λ=1.3 - 1.6μm are integrated. (It should be noted that GaAs-based waveguides will be suitable for λ= 1.3 - 1.55μm if sources and detectors do not need to be integrated monolithically.) Waveguides have been made in InP-based material (25-28) but their performance is not yet as good as GaAs waveguides, because the materials technology is not as well advanced.

4.2 Other non-switchable waveguides

As well as straight waveguides, other elements of integrated optical circuits have been demonstrated in semiconductors, including Y-junctions, polarisers and curved waveguides.

Y-junctions made with InGaAsP buried guides (28) have shown excess splitting losses varying between 0.2dB and 3.3dB as the intersection angle increased from 0.6° to 3°. Y junctions have also been made with n/n^+ InP rib waveguides using chemical etching (29). For an intersection angle of 1° the excess loss was 1dB. This is probably an indication of the extra scattering caused by the rib waveguide. These devices form an essential part of one type of Mach-Zehnder switch, and to reduce further this branching loss, it will be necessary to pay careful attention to the exact shape of the intersection.

Polarisers can be made by loading a waveguide with a metal, when TM modes will absorbed preferentially. A 35dB extinction ratio has been obtained in a 250μm long rib waveguide structure (30).

The performance of curved waveguides is an area in which the use of semiconductors is particularly beneficial. Because guided modes radiate energy from a curved waveguide, tight confinement in both lateral and vertical directions is required. Using a $Ga_{0.88}Al_{0.12}As/GaAs$ waveguide with a relatively high mesa, a loss of 0.6dB/radian has been measured on a curved guide with a 300μm radius of curvature (31). This loss value was thought to have been limited by scattering from edge roughness, rather than radiation due to the bend. The benefits of the strong guiding offered by semiconductors has also been demonstrated by measurements on abrupt bends (28). Using buried InGaAsP/InP waveguides, excess bend losses varied from 0.1dB at an angle of 0.5° to 4.2dB at 3.4°. Because of the relatively small index change induced by Ti indiffusion in $LiNbO_3$, small radii curves and abrupt bends with such low loss are not possible. Curves with a small radius of curvature will be important in achieving high packing densities in optical integrated circuits.

4.3 Switchable devices

Switchable semiconductor devices (using the electro-optic effect) include phase modulators (32), directional couplers (6,9,25), Mach-Zehnder interferometer switches (33) and TE-TM mode converters (13). An attempt to summarise the results on recent semiconductor devices is given in table 1, which also includes some state-of-the-art $LiNbO_3$ devices. It can be extremely difficult to make objective comparisons between different devices - not least because there are often trade-offs between various aspects of performance, and sometimes the less favourable parameters are not discussed in the published literature. The figure of merit that has been taken is the voltage for a π phase shift x length product normalised to wavelength ($V_\pi L/\lambda^2$). The λ^2 term is introduced because the amount of phase shift depends on wavelength (eqn (i)); and the distance over which the electric field is applied is dependent on the mode size, which is also approximately proportional to λ.

Examination of table 1 indicates that the $V_\pi L/\lambda^2$ figure of merit for $LiNbO_3$ and semiconductors are not as dissimilar as the 6:1 difference in n^3r might suggest. This is because the optical/electric field overlap can be higher for semiconductors. However, there is still considerable scope for improvement, as the best semiconductor device (32) is not fully optimised. It is clear from table 1 that semiconductor losses are still significantly higher than those of $LiNbO_3$, and this is the area in which most effort must be directed in the future. The results of refs 34 and 36 indicate the trade-offs between various parameters for $LiNbO_3$ devices, in particular between insertion loss and V_π. With the increased control over electrical and optical fields offered by semiconductor technology, it is hoped that this trade-off will be more favourable in semiconductors.

4 PERFORMANCE OF SEMICONDUCTOR WAVEGUIDE DEVICES

4.1 Straight waveguides

Many studies on waveguides in III-V semiconductors - mainly GaAs - have been reported in the past 12 years, and have encompassed a wide variety of structures. Much of this early work concentrated on using carrier concentration variations to produce an appropriate refractive index profile (18-22). In some cases (18,20), localised diffusion or implantation has produced guiding in both vertical and lateral directions. Although this has advantages in ease of fabrication, it does not allow the independent control of vertical and lateral confinement which is very important for optimising guide performance. As mentioned earlier, the other disadvantage of using carrier concentration variations is the possibility of high absorption loss. The refractive index is only a weak function of carrier density in GaAs and quite high substrate doping densities need to be used ($\approx 10^{18} cm^{-3}$). However,using relatively thick (3.5μm) n^- epilayers and improved fabrication techniques, propagation losses as low as 2dB/cm have recently been measured in n^-/n^+ GaAs waveguides (6). These thick epilayers result in a rather high operating voltage for electro-optically controlled devices. It seems unlikely that devices using carrier concentration variations will give the ultimate in low loss, and they only allow limited flexibility in device design.

With gradually improving control over epitaxial growth techniques it has become possible to make low loss waveguides using compositional variations to produce the refractive index profile. Single mode GaAs/GaAlAs waveguide structures have been demonstrated using LPE (9) with propagation losses as low as 1dB/cm at λ=1.56μm and similar structures made with MBE-grown material (7) have given losses <2dB/cm at λ=1.15μm. MO-CVD has also been used, and multimode waveguides using a GaAs/GaAlAs structure have given losses of 1.4dB/cm at λ=1.3μm (8). It seems likely that this approach of using compositional variations to provide the necessary refractive index profiles is the best approach for low propagation losses, and allows most flexibility in device design.

The residual losses between 1 and 2dB/cm measured on these heterostructure waveguides seem likely to be due to scattering from the waveguide edges. Improvements in waveguide fabrication technology should improve this. Several methods for etching semiconductor rib waveguides have been reported, including wet chemical etching (8), argon ion beam milling (7) and reactive ion etching (6), each with their own advantages and disadvantages. Fig 3a and b show waveguides produced by wet chemical etching and ion beam milling respectively. Both patterns were replicated using a photoresist mask.

(a) (b)

Fig 3. (a) InP waveguide made by chemical etching; stripe is 5μm wide and 1.5μm high.
(b) GaAs/GaAlAs waveguide made by ion beam milling; stripe is 5μm wide and 1.1μm high

A significant feature of the chemically etched waveguide is that the chemical used was selective to a specific crystallographic direction, and thus extremely smooth guide walls were obtained. However, one of the disadvantages of chemical etching is that it is not possible to maintain this mesa shape around a curve, - and curves are particularly important in many waveguide devices. The ion-beam milling technique is a very well controlled process giving high yields. Curved guides can be formed readily with no change in mesa shape with direction. However, there is a certain edge roughness which is thought to be associated with the photoresist pattern. Clearly there is considerable scope to improve rib roughness - possibly a combination of dry and wet etching might be optimum.

Besides these rib waveguides using planar epitaxial material, some novel approaches have also been pursued. A single mode GaAs waveguide with an oxide confining layer has been made by a technique of lateral epitaxial growth by VPE (23). This showed a loss of 2.3dB/cm. An n/n^+ GaAs waveguide has also been made by localised epitaxy with a propagation loss of 1.5dB/cm (24). The unconventional prismatic shapes of these latter guides may limit their applications, especially if electric fields need to be applied.

semi-insulating semiconductor substrate material, - with this it might be possible to use similar techniques to those used for insulators.

3.3.3 Speed

For conventional lumped-contact waveguide devices, the speed is limited by the RC time constant. Two factors favour semiconductors for high speed operation. First, the dielectric constant for GaAs is 5 times less than that for $LiNbO_3$ resulting in lower capacitance for similar geometries, and secondly, the capacitance of semiconductor devices is determined by the thickness of the space-charge region. This can be made very thick, if necessary, with little voltage penalty using doping profiles similar to those used in advanced III-V avalanche photodiodes (14). Thus it should be possible to reach the photon transit-time limit of 113ps/cm without compromising other device parameters.

The ultimate speed will be obtained by using a travelling-wave (TW) electrode configuration. The performance of $LiNbO_3$ TW devices is limited by the differing electrical and optical propagation speeds. This can be overcome to a certain extent by using special electrode configurations (15) but this results in a band-pass characteristic which can be undesirable in some applications. In GaAs the dielectric constant is approximately equal to the refractive index squared so, in principle, it should be possible to achieve much higher operating speeds over a very broad bandwidth. An essential requirement for semiconductor TW devices is high quality, low loss substrate material. To date, no travelling wave devices made in semiconductors have been reported.

3.3.4 Acousto-optics

III-V semiconductors have not been favoured for acousto-optic applications, and this might seem surprising in view of the fact that the acousto-optic figure of merit for GaAs is 15 times greater than that of $LiNbO_3$, one of the most popular acousto-optic materials. The two problems with III-V semiconductors are that i) the piezo-electric effect is very small, making generation of the acoustic waves difficult, and ii) the propagation loss of the acoustic waves in the semiconductor is very high. The generation of acoustic waves can be achieved by using a ZnO transducer (16), but on the present evidence it would seem that III-V semiconductors are unlikely to compete with $LiNbO_3$ for acousto-optic applications.

3.3.5 Stability, reliability, fabrication and cost

A particular problem experienced with $LiNbO_3$ waveguides is the photorefractive effect - a permanent change in the refractive index profile induced by the transmitted optical flux. This has meant that optical power levels have to be kept low, and has resulted in dynamic range limitations in spectrum analysers, and other difficulties. These problems have been alleviated to a certain extent by using longer wavelengths (1). However, a full understanding of this behaviour and complete elimination of these undesirable effects awaits further detailed study. An additional problem that can occur, is ionic drift under dc operation. The remedy here is probably higher quality buffer layers (usually SiO_2), and again, further work is required. Another difficulty with $LiNbO_3$ is the sensitivity of some properties (notably birefringence) to temperature.

Clearly semiconductor devices have not been investigated in as much detail as $LiNbO_3$ devices, but to date no particular reliability or stability problems have been identified. One advantage here is that there is considerable experience in assessing, understanding and rectifying reliability problems with other III-V semiconductor devices, and this should be useful in speeding the progress of research-type devices towards commercial viability.

Cost may be of considerable importance in some cases - the requirement for high capacity telecommunications links in the local network could be satisfied by a fully coherent transmission system (48) rather than the conventional direct detection systems, and this would entail the use of a large number of waveguide devices. It is extremely difficult to estimate costs for waveguide devices, because, amongst other factors, it will depend on the complexity of the device and the numbers required. However, unlike the case of silicon devices, the cost of the starting material may be significant, particularly for the larger area devices. At present the cost of GaAs material is between £15 and £35 per square inch, and InP between £60 and £70 per square inch. These costs are likely to reduce because of the increasing number of applications for III-V semiconductor devices (lasers, detectors, electronic ICs &c), and may well approach that of $LiNbO_3$ (£8 - £25 per square inch). However, a more serious problem is the availability of sufficiently large area III-V wafers. Because, in general, throughput of III-V semiconductor devices is relatively small at present, the motivation to increase slice areas from the standard 2" diameter is limited. This could present a serious limitation for the fabrication of large area semiconductor optical integrated circuits. It is hoped that larger area, high quality slices will emerge with increasing requirements for larger scale GaAs electronic integrated circuits

At the moment, the fabrication of $LiNbO_3$ devices is relatively easy compared with semiconductor devices. It is the epitaxial growth of the semiconductor layers that is the major problem; apart from this and the chemical etching of rib waveguides, much of the rest of the fabrication is similar to that of $LiNbO_3$ (17). However, the ability to cleave semiconductors does confer certain advantages, and accurate polishing of the facets is not required. If MO-VPE and/or MBE can realise their potential as high yield growth processes, it may be that the fabrication of semiconductor waveguide devices will eventually not be significantly more expensive than making $LiNbO_3$ devices.

on bulk GaAs (4) give a loss of about 13dB/cm at a wavelength between 1 and 1.5μm for an electron density of $5\times10^{17}cm^{-3}$. This explains why so many of the devices made with an n^+ substrate (typically $>10^{18}cm^{-3}$) near the guiding region show such high losses. Bulk InP shows a lower loss - 7dB/cm at $10^{18}cm^{-3}$ at a wavelength of 1.3μm (5) - due to its different band parameters, but the best way of decreasing free carrier absorption losses is to reduce the number of free carriers. There have been no published reports of free-carrier absorption values in bulk material with low doping in the wavelength region of interest, but by extrapolation from published figures the free-carrier absorption is expected to be negligible at electron concentrations less than $10^{16}cm^{-3}$ (4,5). Hole concentrations will need to be much lower because of inter-valence band transitions.

Recently, propagation losses between 1 and 2dB/cm have been reported for devices with relatively low doping levels (6-9), and it is thought that these values may be limited by scattering from the rough edges of the rib structure that is used to confine the radiation in the lateral direction. Improved methods of rib fabrication, possibly involving such techniques as reactive-ion etching, the use of planar waveguide structures and the growth of epitaxial layers over a rib structure should reduce this scattering loss.

There is still some way to go to achieve acceptably low propagation losses. (0.5dB/cm might be acceptable initially for discrete devices, but losses may have to be less than this for larger scale integrated circuits.) The indications to date are that there are no inherent barriers to achieving these figures, and that with improved epitaxial growth (notably lower doping levels), and better fabrication techniques (smoother rib waveguides, buried waveguides) the values demonstrated in $LiNbO_3$ (0.3dB/cm for single mode waveguides at 1.3μm; 0.1dB/cm in bulk material at 1.15μm (10)) should be realised in semiconductors in the near future.

Coupling radiation from an optical fibre or laser into a waveguide device (or vice-versa) is limited by Fresnel reflection losses and the mismatch of the optical fields. The former limitation can be effectively eliminated by using λ/4 thickness anti-reflection coatings. The latter may be alleviated by modifying the waveguide refractive index profile. In a semiconductor rib waveguide several parameters can be adjusted - eg waveguide composition, layer thicknesses, rib height, rib width. However, it is clear that the $V_\pi L$ product may be compromised slightly as the guiding region thickness is varied. No systematic study on this aspect has been carried out yet, but a loss of 1dB has been measured for butt-joint coupling of an n^-/n^+ GaAs rib waveguide with a single mode fibre (6). Although not as good as some of the results for coupling fibres to $LiNbO_3$ devices, it is a very encouraging result, and further improvements are expected. Of course, many of these coupling problems will be considerably reduced when lasers and detectors are integrated monolithically with the waveguide devices.

3.3 Other differences

Although the electro-optic coefficient and loss are generally considered the most significant differences between semiconductors and $LiNbO_3$, there are a number of other properties of relevance to waveguide devices. They will now be considered briefly.

3.3.1 Birefringence

$LiNbO_3$ has a high natural birefringence ($\Delta N=0.1$). Although this can be undesirable in some instances, it can also be used to good advantage in devices using TE-TM mode conversion. In such devices, phase matching is required for coupling between the orthogonal polarisations, and this is usually achieved by some sort of periodicity in the electrode configuration. This gives a wavelength dependence which can form the basis of a filter (11). However, phase matching in $LiNbO_3$ usually results in rather narrow bandwidth operation (3.6nm for the filter mentioned above) which will be undesirable for some applications.

III-V semiconductors on the other hand are optically isotropic and only a small birefringence is introduced when making a waveguide structure. However, it has been demonstrated that birefringence can be induced artificially by using a large number of layers of different refractive index (12). A birefringence of $\Delta N=0.055$ has been demonstrated using a GaAs/AlAs multilayer structure. Although this approach has not been pursued for device applications, it seems that this method could give great flexibility if semiconductor devices with a controllable amount of birefringence are required.

3.3.2 Electrical properties

$LiNbO_3$ is an insulator, and this confers particular advantages when making composite components - ie integrated structures incorporating a variety of different devices. For a particular substrate orientation, different electro-optic tensor elements are needed to change the propagation constants of TE and TM modes, and yet another one to control the amount of TE-TM mode conversion. These different elements can be used by applying electric fields in different directions. An insulator is ideally suited to this approach because independent electric fields may be applied on different parts of the slice in a variety of directions.

In a semiconductor, an electric field is applied via a Schottky barrier or pn junction, and applying fields in different directions is much more difficult. In fact, to date, the only way to utilise off-diagonal tensor elements has been to use differently orientated substrates. The normally used substrate is (001) with propagation in the <110> direction. Applying an electric field in the <001> direction changes the propagation constant of the TE mode. TE-TM mode conversion has been demonstrated by using (110) substrates with an electric field in the <110> direction (13). However, in principle it is possible to devise electrode schemes which will allow the flexibility offered by insulators (52), but the fabrication of such devices will be considerably more complex. An alternative approach would be to use high quality

between optical and electrical fields, and it should be minimised for low voltage operation.

The overlap factor Γ can be controlled to a limited degree in $LiNbO_3$ by adjusting electrode separations and varying the guide refractive index profile to optimise the overlap. Values of Γ less than 0.5 are normally obtained, and values can be considerably less than this if the device is required to have high speed and/or low coupling loss.

With semiconductors there is more scope to optimise Γ because the optical and electric field profiles may be controlled independently. The optical field distribution is determined by compositional changes, whilst the electric field distribution depends on the doping profile. This point can be illustrated with reference to fig 2.

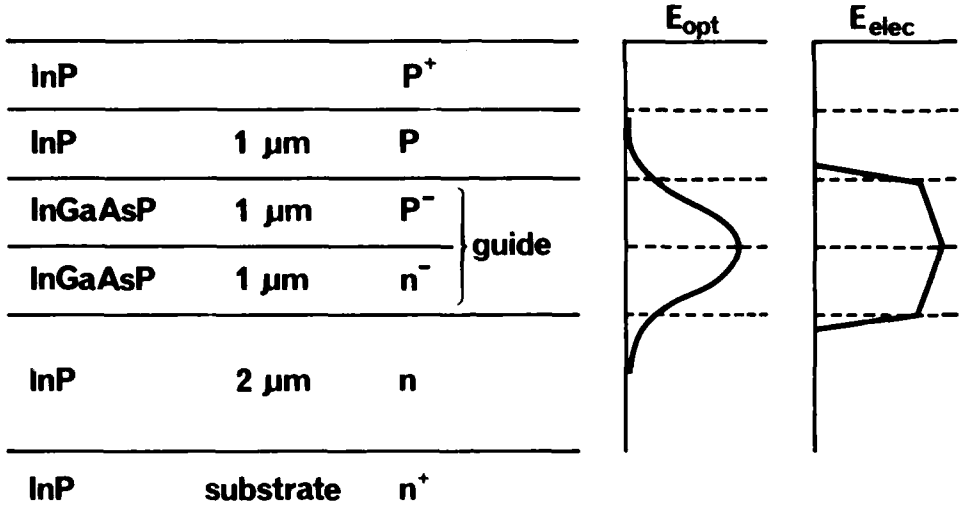


FIG 2 Possible slab waveguide structure

The waveguiding region consists of a p^-/n^- junction of $In_{1-x}Ga_xAs_yP_{1-y}$. For a 2μm thick guiding region the value of the refractive index step (ΔN) needed for monomode guiding in this symmetric waveguide structure is about .006 at a wavelength of 1.3μm. For $InP/In_{1-x}Ga_xAs_yP_{1-y}$ $\Delta N \approx 0.3y$ and as $y \approx 2.2x$ for lattice-matching to InP, the composition of the guiding layers should be $In_{0.99}Ga_{0.01}As_{0.02}P_{0.98}$. (It is assumed that the doping levels are sufficiently low that the free carrier contribution to the refractive index can be ignored.) Adjustments to the layer thickness and composition could give optimised coupling. The electric field profile however, is determined by the applied voltage and the doping profile. In the case that $p \approx 10p^-$ and $n \approx 10n^-$, most of the voltage will be dropped over the guiding region, and the electric field profile will be as in fig 2. If however the doping varied gradually through the structure it could, in principle, be possible to engineer a perfect overlap between the electric and optical fields. It is envisaged that the ability to vary the doping in this manner will become available with the development of MO-VPE and MBE. Semiconductor technology is not yet sufficiently advanced to be able to make such optimised structures, but in section 4 it will be mentioned that the use of double heterostructure waveguides with appropriate doping profiles has already enabled the demonstration of semiconductor devices with $V_\pi L$ products comparable with those of good $LiNbO_3$ devices.

3.2 Loss

The transmission loss of a waveguide can be extremely important, especially in applications such as telecommunications where optical power is at a premium. For other applications loss may not be too much of a problem. However, it is clear that the high losses shown by semiconductors to date have been viewed as a major disadvantage.

GaAs/GaAlAs and InP/InGaAsP-based waveguide devices cannot be used for wavelengths $<0.8\mu m$ because of band-edge absorption. It may be possible to extend this wavelength range in the future by using other III-V semiconductors, but at present it is more realistic to consider semiconductors solely for applications with $\lambda > 0.85\mu m$.

Transmission loss can be separated into two components - propagation losses and coupling loss. In semiconductor waveguides the wavelength of the propagating radiation usually corresponds to a photon energy considerably less than the bandgap energy. Under these circumstances, free-carrier absorption is expected to be the dominant absorption mechanism. This includes inter and intra-conduction and valence band transitions. The free-carrier absorption coefficient is directly proportional to the number of free carriers. Measurements

continuing advances in the field, and this paper concludes with some comments on the developments that will be needed.

2 PRINCIPLES OF SEMICONDUCTOR WAVEGUIDES

A variety of semiconductor waveguide structures have been demonstrated. Some of these will be discussed later in the paper, but initially the essential elements of a semiconductor waveguide will be discussed with respect to a rib waveguide; probably the most commonly used structure (fig 1).

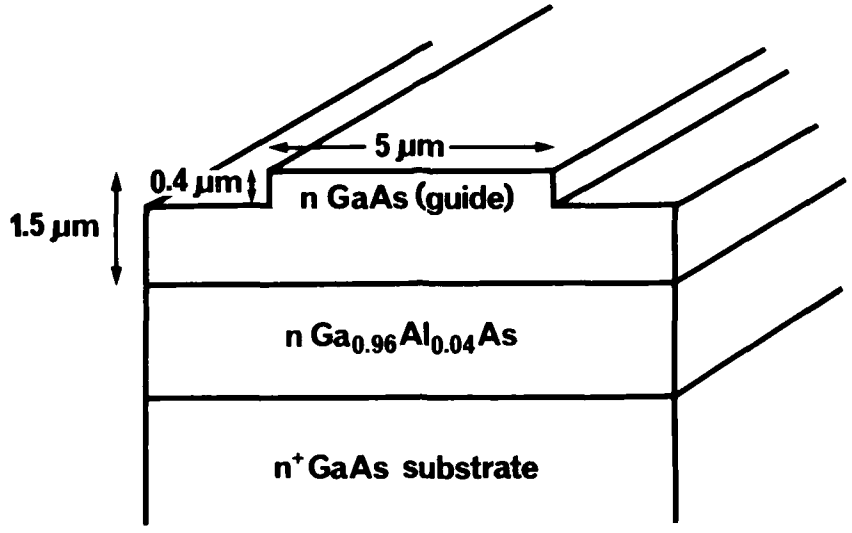


FIG 1 A GaAs/GaAlAs Rib Waveguide

As in any waveguide structure, an appropriate refractive index profile is required. In fig 1 the guiding layer (GaAs) is bounded in the vertical direction by a layer of $Ga_{0.96}Al_{0.04}As$ and air. The refractive index difference between GaAs and $Ga_{1-x}Al_xAs$ is approximately 0.4x. The dimensions shown in fig 1 are typical for monomode operation.

In the lateral direction, the confinement is obtained by etching a rib structure in the semiconductor. In many cases it will be necessary to apply an electric field to the guiding region, and this can be done with a Schottky contact, obtained by having an appropriate metal layer on top of the rib. An ohmic contact would be applied at the base of the substrate.

3 THE PROPERTIES OF SEMICONDUCTORS RELEVANT TO WAVEGUIDE DEVICES

3.1 Electro-optic coefficient

The figure of merit for many waveguide devices is related to the change in propagation constant that can be induced by an electric field, for example:-

$$\Delta\beta = \pi n^3 r E' L/\lambda \qquad \ldots\ldots(i)$$

where n is the effective refractive index of the waveguide for the propagating mode, r is the electro-optic tensor element appropriate to the crystal orientation and the direction of the applied electric field and L is the device length. E' is an 'effective' electric field, determined by the applied voltage (V), the separation of the electrodes (d) and a factor (Γ) giving the overlap between the optical (E_{opt}) and electric (E_{elec}) fields.

The figure of merit most often quoted is n^3r. $LiNbO_3$ belongs to the 3m point-group symmetry, and has four non-identical electro-optic coefficients. By using a suitably orientated substrate and electric field direction, the optimum element (r_{33}) can be used with $n^3r_{33}=328\times10^{-12}$m/V. The other elements, r_{22},r_{13} and r_{51} give smaller figures of merit. For GaAs, which has only one coefficient (r_{41}), the figure of merit is about 6 times less ($n^3r_{41}=60\times10^{-12}$m/V) and InP shows a similar value (3). There seems no particular reason to suppose that related III-V ternary and quaternary compounds should be significantly different.

A parameter which is perhaps more useful than n^3r is the voltage length product for a π phase shift ($V_\pi L$). Taking $E'=V\Gamma/d$, then

$$V_\pi L = \lambda d/n^3 r\Gamma \qquad \ldots\ldots(ii)$$

This is a particularly relevant parameter because it incorporates the degree of overlap

THE POTENTIAL OF SEMICONDUCTORS FOR OPTICAL INTEGRATED CIRCUITS

S RITCHIE AND A G STEVENTON

BRITISH TELECOM RESEARCH LABORATORIES
MARTLESHAM HEATH
IPSWICH, IP5 7RE, UK

SUMMARY

Optical circuits made with semiconductors should eventually have considerable performance, reliability and cost benefits over similar components made with many alternative materials, because of the ability to integrate waveguide devices with optical sources, detectors and electronic circuitry. Although, at present, the performance of semiconductor waveguide devices is inferior to that of devices made with lithium niobate, progress is rapid and there seem no fundamental reasons why semiconductor devices should not eventually give performance comparable with many non-semiconductor devices. A wide range of semiconductor optical and electronic devices will be available for integration in semiconductor optical integrated circuits, but considerable technological developments are needed - especially in epitaxial growth.

1 INTRODUCTION

The advantages of using optics for certain signal processing applications are well established. Perhaps the best known example is rf spectrum analysis (1) which can be performed at much higher speeds using optics than electronic methods. Another example where optics could offer much higher speeds is in analogue-to-digital conversion (2). In addition, with the increasing use of optics in telecommunications, both for long and short haul applications, more and more optical processing functions are being required. For example, the more advanced optical systems in the future may require amplitude, phase and frequency modulation, wavelength and time division multiplexing and space switching.

At present the processing of optical signals, either using discrete or integrated components, is usually carried out with dielectric materials, most commonly lithium niobate ($LiNbO_3$). This has the particular advantage that suitable starting material is available, and the fabrication of waveguides for optical devices is relatively easy - in general a titanium diffusion is all that is needed to produce the necessary refractive index profiles. Acousto-optic and electro-optic coefficients are reasonably high and the optical propagation loss of $LiNbO_3$ is sufficiently low that, combined with relatively well developed techniques for coupling to and from optical fibres, acceptably low insertion loss can be achieved.

III-V semiconductors - typically gallium arsenide (GaAs) and indium phosphide (InP) - also exhibit acousto-optic and electro-optic effects and can perform many of the optical processing functions demonstrated in dielectrics. For reasons to be discussed later, III-V semiconductors have not been used widely for acousto-optic applications, and the majority of work on semiconductor waveguide devices has employed the electro-optic effect. The main attraction of using semiconductors is the potential for integrating optical sources, detectors and any electronic components that are necessary on the same substrate, with all the benefits in performance, size, reliability and cost that this would entail. The particular disadvantages of semiconductors most often quoted are that:-

a) they have a relatively low electro-optic coefficient compared with $LiNbO_3$,
b) their measured optical propagation losses have been considerably higher than those in $LiNbO_3$,
c) the growth of suitable epitaxial layers is difficult, and requires expensive equipment.

These, and other, differences between semiconductors and $LiNbO_3$ for waveguide devices will be discussed in more detail in section 3.

A range of semiconductor waveguide devices have been demonstrated and will be described in section 4. These are mainly discrete devices (eg directional coupler switches, TE-TM mode converters &c) and semiconductor components of the complexity of some of the $LiNbO_3$ devices now being made have not been demonstrated yet. However, recent improvements in semiconductor waveguide performance suggest that it is now feasible to consider the fabrication of larger scale structures.

With the development of advanced epitaxial growth techniques such as metallorganic vapour phase epitaxy (MO-VPE) and molecular beam epitaxy (MBE), a variety of novel optical and electronic devices are being realised. The concept of "bandgap engineering" is extending the possibilities for semiconductor device functions, and conventionally accepted limitations to semiconductor device performance are constantly being eroded. Whilst many of these developments are at a very early stage - for discrete devices, let alone integrated structures - the huge potential for semiconductor optoelectronics cannot be ignored. Some of the most interesting of recent developments in III-V semiconductor devices are also discussed in this paper, as well as the status of the monolithic integration of some of the relatively simple structures. The development of semiconductor growth and processing technology is crucial to

DISCUSSION

C.Klingshirn, Ge

If you recombine the two branches in a state such that the transmission is zero then the light is not absorbed — it is just scattered out of this waveguide line, and it goes into the bulk. So if you have one crystal grown on one substrate and many of these structures, is there not a danger of having a lot of crosstalk because the light goes into the bulk and is reflected from the back surface and goes through the whole substrate and comes out wherever it wants.

Author's Reply

That could be a problem -- you're right — but we have some ways to fight against it. We can depolish the under surface or put absorbers or not have parallel surfaces at the bottom and at the top. There are many ways to fight against these problems. The best way may be to make the surface absorbing as can be done for lithium niobate by heating it without oxygen.

approach, the electronic devices are incorporated adjacent to the laser stripe within the cleaved facets. For example, Bar-Chaim et al (93) have integrated a photodiode, a FET and a laser on a semi-insulating GaAs substrate to produce a rudimentary optical repeater with a gain bandwidth product of 178MHz, and they predict improvements to better than 1GHz. Much of this type of work has employed LPE, but larger circuits will need the improved control of MBE and/or MO-VPE, and improved circuit planarity. Sanada et al (94) used MBE on channelled substrates to obtain nearly planar laser/FET circuits. They also used MQW active layers to reduce chip thermal dissipation, and achieved a convertion ratio of laser output to FET gate voltage of 3.3mW/V with rise and fall times of 1ns.

The restrictions imposed by working within the confines of the laser facets are removed by using integrable lasers, and both Matsueda et al (95) and Carney et al (96) describe optoelectronic integrated circuits using such lasers grown in recessed parts of the substrate, in order to improve the planarity for subsequent GaAs IC processing steps. Carney et al describe their development of a transmitter consisting of a laser, drive circuit and 4x1 multiplexer capable of 1Gbit/s operation, although, at that time, the circuit was not fully operational.

The integration of a PIN photodiode and high performance FET is very useful, and progress on both GaAs/GaAs and InGaAs/InP devices have been made (97,98). In the case of the GaAs receivers, a useful transimpedance preamplifier has been integrated with the PIN diode (99). It is clear that quite complex GaAs-based optoelectronic circuits are becoming available at the research stage already, and that the preliminary steps for the 1.3-1.6μm region are progressing rapidly.

Fig 5 shows an example of the optoelectronic integrated circuit (OEIC) concept (100). This particular component would take a number of optical inputs, convert the signals to electrical form, perform various switching operations before reconverting to optical signals again. It is envisaged that such a device will be extremely useful for optical communications, especially in local area networks, and local network distribution schemes. Of course this is not a true optical circuit, as all the processing occurs in electronic form, but with suitable developments in technology the switching could be done in optical form, which would confer advantages for some applications.

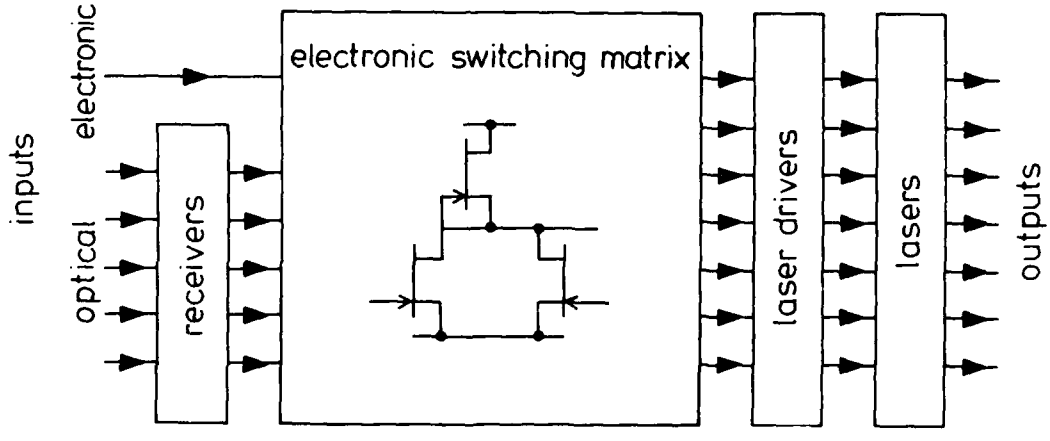


FIG 5 An optoelectronic integrated circuit (OEIC)

The rapid progress in integration of electronic circuits with lasers and detectors has been due principally to the use of improved epitaxial growth methods for depositing the optoelectronic device layers. Progress has clearly been most rapid in GaAs-based circuits where the existing GaAs IC technology may be used directly. However, the benefits of integration for InP-based optoelectronic integrated circuits are large, and an increasing number of laboratories are establishing the techniques needed. Progress on integrating waveguides on the other hand has been slow, partly because of the availability of $LiNbO_3$ devices as an alternative to semiconductors, but also because of lack of suitable material. Now that this is becoming available, and as improved discrete devices are being demonstrated, the waveguide integration is likely to accelerate. The use of waveguide techniques in coherent communications applications would result in highly practical and high performance systems.

9 SEMICONDUCTOR GROWTH AND PROCESSING TECHNOLOGY

The development of suitable epitaxial growth techniques is crucial for the realisation of the advanced electronic and optical devices mentioned in this paper. The most widely used epitaxial technique for III-V semiconductors is liquid phase epitaxy (LPE), but it is generally held that this will not be the most suitable for waveguides, and the more advanced optical and electrical devices. This is primarily because a large area of material with a sufficiently low defect density is hard to obtain. However, some of the best waveguide devices to date have been achieved using LPE (6,9,13).

The techniques on which hopes are based are MO-VPE and MBE. In principle, these techniques should give large areas of material with well controlled composition and thickness. Both MO-VPE and MBE have demonstrated good GaAs/GaAlAs laser structures. Waveguides have also been made out of GaAs-based material grown by these techniques, but further development is needed before optimum structures can be obtained routinely. One limitation at present seems to be doping control - in particular the demonstration of low-doped ($<10^{15}cm^{-3}$) GaAlAs layers. The growth of InP-based materials is at a very rudimentary stage. Although there is considerable work world-wide on adapting the MO-VPE and MBE techniques to InGaAsP and InGaAlAs, LPE is still the only option at present for InP-based heterostructure devices.

Thus it seems that in the immediate future, LPE is likely to be used widely to demonstrate discrete waveguides and the simplest waveguide structures, but that over the next few years MO-VPE and MBE will emerge as the most suitable techniques, allowing much larger scale integration.

In the demonstrations to date of electronic/optical and optical/optical integrated devices, non-optimum epitaxial layer structures have been employed. For example, a heterojunction bipolar transistor has a similar layer structure to that of a semiconductor laser, so that by compromising the performance of both devices, a single layer structure will allow the monolithic integration of both devices. This is inherently unsatisfactory, as the resulting performance can be so inferior as to negate the advantages of integration. In future it will be desirable to grow optimised structures on different parts of the substrate. This could be done by localised area epitaxy. For example, MBE is ideally suited for in-situ shadow masking. For MO-VPE, it may be necessary to use dielectric masking. Alternatively, the use of an epitaxial photodeposition technique may give improved resolution. In this, an ultra-violet laser is used to stimulate growth, either by localised heating or chemical dissociation.

An important implication of these ideas is the requirement to be able to grow on processed and structured substrates. In many growth techniques for conventional structures, the surface of the substrate is etched away immediately prior to growth to give a good, clean surface for nucleation. This might be unacceptable when growing over certain small features, and there will need to be a much greater understanding of the surfaces of semiconductors, and the requirement for good growth.

Concerning advances in device fabrication, the subject of etching has already been raised. It is likely that semiconductor etching will always be important in the fabrication of integrated optical circuits, although such techniques as epitaxial growth through slots etched in dielectric layers may be envisaged. Dry etching processes are likely to become important because of their uniformity, controllability, and high yield. III-V semiconductors have been etched using ion beam milling, plasma etching, reactive ion etching and reactive ion beam etching, but these techniques need much more development, and a greater understanding of the basic processes involved. The topics of particular importance for integrated optics include uniformity over large areas, anisotropic etching, material selective etching and damage effects.

Lithography is extremely important for waveguide devices. Smooth features (on a $<0.1\mu m$ scale) and the ability to replicate patterns over a large area will be more important than very narrow linwidth features. The relative benefits of electron-beam lithography, and the more conventional photolithography will need to be assessed.

A final point worth considering is the incorporation of non-semiconductor materials on semiconductor substrates. Clearly the relatively small electro- optic coefficient of conventional semiconductors and dielectrics is a disadvantage, and various organic materials with much higher electro-optic coefficients and other non-linearities are being studied (101) and may eventually be deposited on semiconductors to allow the combination of optimum materials.

10 CONCLUSIONS

A detailed comparison between semiconductors and $LiNbO_3$ has indicated that semiconductor waveguide device performance should eventually be comparable with that of $LiNbO_3$ in many respects. In particular, the 6:1 difference in the electro-optic figure of merit is offset by the greater ability to achieve good overlaps between optical and electrical fields. Also, although propagation losses are still higher in semiconductor guides, there have been considerable improvements recently, and there appear no fundamental reasons why comparable losses should not eventually be achieved. $LiNbO_3$ will probably always be superior for acousto-optic applications, but there are grounds for believing that semiconductors will ultimately give higher speed performance.

A considerable advantage for $LiNbO_3$ is the availability of large area starting material and relatively easy fabrication techniques. It remains to be seen whether semiconductor epitaxial growth and processing technology can be developed sufficiently to produce components of comparable cost. Stability and reliability problems may hamper the implementation of $LiNbO_3$ devices for some applications.

The advantages of semiconductors that may outweigh minor performance or cost differentials is the ability to integrate a variety of electronic and optical devices on one chip. The development of epitaxial growth techniques has allowed the demonstration of a wide range of novel devices in III-V semiconductors, some of which have already been integrated together in small scale circuits. So far, most of the attention has focussed on optoelectronic integration, ie lasers, detectors and electronic circuitry. The wider incorporation of waveguides in optical circuits is expected in the near future as growth techniques such as MO-

VPE and MBE reach maturity. Semiconductor picosecond optical sources and detectors have been demonstrated in discrete form, and optical bistability has been observed. The possibilities for integrating these devices are promising, bringing the prospect of all-optical logic much closer.

ACKNOWLEDGEMENTS

The authors acknowledge invaluable discussions with many colleagues at British Telecom Research Laboratories. The paper is published with the permission of the Director of Research, British Telecom.

REFERENCES

1 D Mergerian, E C Malarkey and R P Pautienus, "High dynamic range integrated optical rf spectrum analyser", Proc 4th Int Conf on Integrated Optics and Optical Fibre Communication, Tokyo 1983, Paper 30B3-6, (260-261)

2 See for example (a) H F Taylor, "An optical analog-to-digital converter - design and analysis", IEEE J Quant Electron, Vol QE-15, No 4, 1979, (210-216) and (b) R A Becker and F J Leonberger "2 bit 1Gsample/sec electrooptic guided wave analogue-to-digital converter" IEEE J Quant Electron, Vol QE-8, No 10, 1982, (1411-1413)

3 N Susuki and K Tada, "Electrooptic and acoustooptic properties of InP", Proc 4th Int Conf on Integrated Optics and Optical Fibre Communication, Tokyo 1983, Paper 30B3-1, (250-251)

4 W G Spitzer and J M Whelan, "Infrared absorption and electron effective mass in n-type gallium arsenide", Phys Rev, Vol 114, No 1, 1959, (59-63)

5 O K Kim and W A Bonner, "Infrared reflectance and absorption of n-type InP", J Electronic Mats, Vol 12, No 5, 1983, (827-836)

6 P Buchmann, H Kaufmann, H Melchior and G Guekos, "Reactive ion etched GaAs optical waveguide modulators with low loss and high speed", Electron Lett, Vol 20, No 7, 1984, (295-297)

7 A J N Houghton, D A Andrews, G J Davies and S Ritchie, "Low-loss optical waveguides in MBE-grown GaAs/GaAlAs heterostructures", Optics Communications, Vol 46, No 3/4, 1983, (164-166)

8 R G Walker and R C Goodfellow, "Attenuation measurements on MO-CVD-grown GaAs/GaAlAs optical waveguides", Electron Lett, Vol 19, No 15, 1983, (590-592)

9 A Carenco, P Sansonetti, L Menigaux, J Brandon and M Rondot, "Low loss GaAs-Al_x $Ga_{1-x}As$ double heterostructure directional coupler", Proc 7th Topical Meeting on Integrated and Guided Wave Optics, Kissimmee, 1984, (Paper ThB4)

10 R C Alferness, V R Ramaswamy, S K Korotky, M D Divine and L L Buhl, "Efficient single-mode fiber to titanium diffused lithium niobate waveguide coupling for λ=1.32μm", IEEE J Quant Electron, Vol QE-18, No 10, 1982, (1807-1813)

11 R C Alferness and L L Buhl, "Long wavelength Ti:$LiNbO_3$ waveguide electro-optic TE-TM converter", Electron Lett, Vol 19, No 2, 1983, (40-41)

12 J P van der Ziel, M Illegems and R M Mikulyak, "Optical birefringence of thin GaAs-GaAlAs multilayer films", Appl Phys Lett, Vol 28, 1976, (735)

13 F K Reinhart, R A Logan and W R Sinclair, "Electrooptic polarisation modulation in multielectrode Al_x $Ga_{1-x}As$ rib waveguides", IEEE J Quant Electron, Vol QE-18, No 4, 1982, (763-766)

14 S R Forrest, R G Smith and O K Kim, "Performance of $In_{0.53}Ga_{0.47}As/InP$ avalanche photodiodes", IEEE J Quant Electron, Vol QE-18, No 12, 1982, (2040-2048)

15 R C Alferness, S K Korotky and E A J Marcatili, "Velocity-matching techniques for integrated optic traveling wave switch/modulators", IEEE J Quant Electron, Vol QE-20, No 3, 1984, (301-309)

16 C J Lii, C C Lee, O Yamazaki and L S Yap, "Efficient wideband acoustooptic Bragg diffraction in GaAs-GaAlAs waveguide structures", Proc 4th Int Conf on Integrated Optics and Optical Fibre Communication, Tokyo 1983, Paper 30B3-2, (252-253)

17 A J N Houghton and P M Rodgers, "The fabrication of guided wave optical modulators and switches in semiconductors", Br Telecom Technol J, Vol 1, No 2, 1983, (78-81)

18 E Garmire, D F Lovelace and G H B Thompson, "Diffused two-dimensional optical waveguides in GaAs", Appl Phys Lett, Vol 26, No 6, 1975, (329-331)

19 F J Leonberger, J P Donnelly and C O Bozler, "Low-loss GaAs $p^+n^-n^+$ three-dimensional optical waveguides", Appl Phys Lett, Vol 28, No 10, 1976, (616-619)

20 S Somekh, E Garmire, A Yariv, H L Garvin and R G Hunsperger, "Channel optical waveguide directional couplers" Appl Phys Lett, Vol 22, No 2, 1973, (46-47)

21 E Garmire, H Stoll, A Yariv and R G Hunsperger, "Optical waveguiding in proton implanted

GaAs", Appl Phys Lett, Vol 21, No 3, 1972, (87-88)

22 F A Blum, D W Shaw and W C Holton, "Optical striplines for integrated optical circuits in epitaxial GaAs", Appl Phys Lett, Vol 25, No 2, 1974, (116-118)

23 F J Leonberger, C O Bozler, R W McClelland and I Melngailis, "Low-loss GaAs optical waveguides formed by lateral epitaxial growth over oxide", Appl Phys Lett, Vol 38, No 5, 1981, (313-315)

24 M Erman, N Vodjdani, J B Theeten and J P Cabanie, "Low loss waveguides grown on GaAs using localised vapour phase epitaxy", Appl Phys Lett, Vol 43, No 10, 1983, (894-895)

25 A Carenco, L Menigaux and N T Linh, "InP electro-optic directional coupler", Appl Phys Lett, Vol 40, No 8, 1982, (653-655)

26 C Bornholdt et al, "Passive optical GaInAsP/InP waveguides", Electron Lett, Vol 19, No 3, 1983, (81-82)

27 O Mikami, H Nakagome and T Saitoh, "GaInAsP/InP buried-heterostructure optical waveguides at 1.5μm wavelength", Electron Lett, Vol 19, No 15, 1983, (593-595)

28 L M Johnson, Z L Liau and S H Groves, "Low-loss GaInAsP buried-heterostructure optical waveguide branches and bends", Appl Phys Lett, Vol 44, No 3, 1984, (278-280)

29 P Buchmann and A J N Houghton, "Optical Y-junctions and S bends formed by preferentially etched single-mode rib waveguides in InP", Electron Lett, Vol 18, No 19, 1982, (850-852)

30 F K Reinhart, J C Shelton, R A Logan and B W Lee, "MOS rib waveguide polarisers", Appl Phys Lett, Vol 36, No 4, 1980, (237-240)

31 M W Austin and P G Flavin, "Small-radii curved waveguides in GaAs/GaAlAs using electron-beam lithography", IEEE J Lightwave Technology, Vol LT-1, No 1, 1983, (236-240)

32 A J N Houghton, P M Rodgers, and D A Andrews, "High performance GaAs/GaAlAs phase modulators for PSK optical fibre systems", Electron Lett, Vol 20, No 11, 1984, (479-481)

33 J P Donnelly, N L DeMeo, G A Ferrante, K B Nichols and F J O'Donnell, "A gallium arsenide electro-optical interferometric modulator", Proc 7th Topical Meeting on Integrated and Guided Wave Optics, Kissimmee, 1984, (Paper ThB3)

34 L Riviere, R Guglielmi and A Carenco, "Single-mode fibre pigtailed 1.5μm fast electro-optic modulator", Proc 4th Int Conf on Integrated Optics and Optical Fibre Communication, Tokyo 1983, (Paper 29C4-4), (362-363)

35 M Kondo, K Komatsu and Y Ohta, "Fibre-coupling loss and drive voltage simultaneous reductions by tapered titanium thickness diffusion on $LiNbO_3$ waveguide switches", Proc 7th Topical Meeting on Integrated and Guided Wave Optics, Kissimmee, 1984, (Paper TuA5)

36 R C Alferness, L L Buhl, S K Korotky and M D Divino, "Ti:$LiNbO_3$ guided wave devices for long wavelength telecommunications", Proc 7th Topical Meeting on Integrated and Guided Wave Optics, Kissimmee, 1984, (Paper TuA1)

37 R A Becker, "Broadband Ti:$LiNbO_3$ guided-wave lumped-element and travelling-wave interferometric modulators", Proc 7th Topical Meeting on Integrated and Guided Wave Optics, Kissimmee, 1984, (Paper TuA2)

38 F J Leonberger "High speed operation of $LiNbO_3$ electrooptic interferometric waveguide modulators", Opt Lett, Vol 5, 1980, (312-314)

39 P Auracher and R Keil, "Method for measuring the rf modulation characteristics of Mach-Zehnder type modulators", Appl Phys Lett, Vol 36, 1980, (626-629)

40 A Yariv and P Yeh, "Optical waves in crystals", Table 12.2, Wiley-Interscience, 1984

41 D A B Miller, "Refractive Fabry-Perot bistability with linear absorption: theory of operation and cavity optimisation", IEEE J Quant Electron, Vol QE-17, No 3, 1981, (306-311)

42 H M Gibbs et al, "Optical bistability in semiconductors", Appl Phys Lett, Vol 35, No 6, 1979, (451-453)

43 H M Gibbs et al, "Room termperature excitonic optical bistability in a GaAs-GaAlAs superlattice etalon", Appl Phys Lett, Vol 41, No 3, 1982, (221-222)

44 C D Poole and E Garmire, "Optical bistability at the bandgap in InAs", Appl Phys Lett, Vol 44, No 4, 1984, (363-365)

45 C T Seaton et al, "Realisation of an InSb bistable device as an optical AND gate and its use to measure recombination times", Appl Phys Lett, Vol 42, No 2, 1983, (131-133)

46 S S Tarng et al, "Use of a diode laser to observe room-temperature, low-power optical bistability in a GaAs-AlGaAs etalon", Appl Phy Lett, Vol 44, No 4, 1984, (360-361)

47 P S Cross, R A Baumgartner and B H Kolner, "A travelling-wave microwave optical modulator", Proc SPIE, Picosecond Optoelectronics, Vol 439, 1983

48 J Stern, "Optical wideband subscriber loops and local area networks in the UK", Proc Int Conf on Communications, May 1984, (884-887)

49 L D Westbrook, P J Fiddyment and P N Robson, "Photoelastic channel optical waveguides in epitaxial GaAs layers", Electron Lett, Vol 16, No5, 1980, (169-170)

50 T M Benson, T Murotani, P N Robson and P A Houston, "A novel electro-optically controlled directional-coupler switch in GaAs epitaxial layers at 1.15μm", IEEE Trans Electron Dev, Vol ED-19, No 9, 1982, (1477-1483)

51 P W Smith, "Hybrid bistable optical devices", Optical Engineering, Vol 19, No 4, 1980, (456-462)

52 B G Kushner, "Two-dimensional electro-optic polarisation modulator for integrated optics", Proc SPIE Vol 408, Integrated Optics III, 1983, (140-144)

53 G J Aspin, J E Carroll and R G Plumb,"The effect of cavity length on picosecond pulse generation with highly rf modulated AlGaAs double heterostructure lasers",Appl Phys Lett, Vol 39, No 11, 1981, (860-861)

54 S Lundqvist, T Andersson and S T Eng,"Generation of tunable single-mode picosecond pulses from an AlGaAs semiconductor laser with grating feedback", Appl Phys Lett, Vol 43, No 8, 1983, (715-717)

55 J M Wiesenfeld and R A Logan, "Broadband tunable picosecond semiconductor lasers", Appl Phys Lett, Vol 39, No 2, 1981, (142-145)

56 D Z Tsang and J N Walpole, "Q-switched semiconductor diode lasers", IEEE J Quant Electron. QE-19, No 2, 1983, (145-156)

57 P T Ho, "Picosecond pulse generation with semiconductor diode lasers", Proc SPIE, Picosecond Optoelectronics, San Diego, 1983, (42-48)

58 D J Bradley, M B Holbrook and W E Sleat,"Bandwidth limited picosecond pulses from an actively-modelocked GaAlAs diode laser", IEEE J Quant Electron, Vol QE-17, 1981, (658-662)

59 J P Van der Ziel, "Generation of sub-picosecond pulses from modelocked AlGaAs semiconductor lasers", Proc SPIE, Picosecond Optoelectronics, San Diego, 1983, (49-55)

60 M Izutsu and T Sueta, "Picosecond pulse response of broad-band guided-wave interferometric light modulators", IEEE J Quant Electron, Vol QE-19, No 4, 1983, (668-673)

61 H A Haus, S T Kirsh, K Mathyssek and P J Leonberger, "Picosecond optical sampling", IEEE J Quant Electron, Vol QE-16, No 8, 1980, (870-874)

62 E A J Marcatili, "Optical subpicosecond gate", Appl Opt, Vol 19, No 9, 1980, (1468-1476)

63 D Z Tsang, J N Walpole, S H Groves, and Z L Liau, "14GHz operation of Q-switched diode lasers", Proc 42nd Dev Research Conf, Santa Barbara, 1984, paper IVA-3

64 C Harder, J S Smith, K Y Lau and A Yariv, "Passive modelocking of of buried heterostructure lasers with nonuniform current injection", Appl Phys Lett, Vol 42, No 9, 1983, (772-774)

65 Y Tada, H Yokoyama, H Ito and H Inaba, "Ultra-short optical pulse generation from microwave modulated AlGaAs diode laser with SELFOC rod resonator" Optics Communications, Vol 47, No 3, 1983, (187-189)

66 S K Korotky, R C Alferness, L L Buhl, C H Joyner and E A J Marcatili, "High- speed pulse generation using a sinusoidally driven Ti:$LiNbO_3$ directional coupler travelling-wave optical modulator", Electron Lett, Vol 20, No 9, 1984, (384-385)

67 L A Molter-Orr, H A Haus, and P J Leonberger, "20 GHz optical waveguide sampler", IEEE J Quant Electron, Vol QE-19, No 12, 1983, (1877-1883)

68 W Roth, H Schumacher and H Beneking, "Fast photoconductive GaAs detectors made by laser stimulated MOCVD", Electron Lett, Vol 19, No 4, 1983, (142-143)

69 R B Hammond, R S Wagner, and N G Paulter,"InP:Fe picosecond photoconductors", Proc SPIE, Picosecond Optoelectronics, San Diego, 1983, (192-196)

70 S Y Wang and D M Bloom, "100 GHz bandwidth planar GaAs Schottky photodiode", Electron Lett, Vol 19, No 14, 1983, (554-555)

71 J M Weisenfeld, A R Chraplyvy, J Stone and C A Burrus, "Measurement of very-high-speed photodetectors with picosecond InGaAsP film lasers" Electron Lett, Vol 19, No 1, 1983, (22-24)

72 N Holonyak, R M Kolban, R D Dupuis, and P D Dapkus, "Quantum-well heterostructure lasers", IEEE J Quant Electron, Vol QE-16, No 2, 1980, (170-184)

73 G C Osbourn, "Strained layer superlattices from lattice mismatched materials", J Appl Phys, Vol 53, No 3, 1982, (1586-1589)

74 G C Osbourn, R M Biefield and P L Gourley, "A GaAsP/GaP strained layer superlattice", Appl Phys Lett, Vol 41, No 2, 1982, (172-174)

75 R D Burnham, W Streifer, D R Scifres, C Lindstrom and T L Paoli,"Low-threshold single quantum well (60A) GaAlAs lasers grown by MO-CVD with Mg as p-type dopant", Electron Lett, Vol 18, No 25, 1982, (1095-1097)

76 Y-H Wu, M Werner and S Wang, "Channeled-substrate GaAs/AlGaAs multiple quantum-well lasers grown by molecular beam epitaxy" Proc 7th Topical Meeting on Integrated and Guided-Wave Optics, Kissimmee, 1984, paper TuC2

77 H Temkin, K Alavi, W R Wagner, T P Pearsall and A Y Cho, "1.5-1.6μm $Ga_{0.47}In_{0.53}As$ multiquantum well lasers grown by molecular beam epitaxy", Appl Phys Lett, Vol 42, No 10, 1983, (845-847)

78 M G Burt, "Gain spectra of quantum-well lasers", Electron Lett, Vol 19, No 6, 1983, (210-211)

79 M G Burt, "Enhanced bandgap resonant nonlinear susceptibility in quantum- well heterostructures", Electron Lett, Vol 19, No 4, 1983, (132-133)

80 M G Burt, "Linewidth enhancement factor for quantum-well lasers", Electron Lett, Vol 20, No 1, 1984, (27-29)

81 C Smith, R A Abram and M G Burt, "Auger recombination in a quantum well heterostructure", J Phys C:Solid State Phys, Vol 16, 1983, (L171-L175)

82 G H Dohler, "n-i-p-i doping superlattices - metastable semiconductors with tunable properties", J Vac Sci Technol, Vol B1, No 2, 1983, (278-284)

83 H Jung, G H Dohler, E O Gobel and K Ploog, "Optical gain in GaAs doping structures" Appl Phys Lett, Vol 43, No 1, 1983, (40-42)

84 R C Hooper, J E Midwinter, D W Smith and I W Stanley, "Progress in monomode transmission techniques in the United Kingdom", J Lightwave Tech, Vol LT-1, No 4, 1983, (596-611)

85 H Morkoc and P M Solomon, "The HEMT: a superfast transistor", IEEE Spectrum, Feb 1984, (28-35)

86 S Adachi, H Kawaguchi, T Takehei and Y Noguchi, "InGaAsP/InP buried-heterostructure lasers (λ=1.5μm) with chemically etched mirrors", J Appl Phys, Vol 52, No 9, 1981, (5843-5845)

87 U Koren, Z Rav-Noy, A Hasson, T R Chen, K L Yu, L C Chiu, S Margalit and A Yariv, "Short cavity InGaAsP/InP lasers with dielectric mirrors", Appl Phys Lett, Vol 42, No 10, 1983, (848-850)

88 N Matsumoto and K Kumabe, "AlGaAs-GaAs semiconductor ring laser", Jap J Appl Phys, Vol 16, No 8, 1977, (1795-1798)

89 L D Westbrook, A W Nelson, P F Fiddyment and J S Evans, "Continuous wave operation of 1.5μm distributed-feedback ridge-waveguide lasers", Electron Lett, Vol 20, No 6, 1984, (225-226)

90 Y Abe et al "Room-temperature cw operation of 1.6μm GaInAsP/InP buried-heterostructure integrated laser with butt-jointed built-in distributed-Bragg-reflection waveguide", Electron Lett, Vol 18, No 10, 1982, (410-411)

91 K Aiki, M Nakamura and J Umeda, "Frequency multiplexing light source with monolithically integrated distributed-feedback diode lasers", Appl Phys Lett, Vol 29, No 8, 1976, (506-508)

92 J L Merz, Y R Yuan and L Perillo, "Monolithic integration of GaAs DH lasers and detectors with curved and channeled-substrate waveguides", Proc 4th Int Conf on Integrated Optics and Optical Fibre Communication, Tokyo, 1983, paper 28B2-1

93 N Bar-Chaim, K Y Lau and I Ury,"Monolithic optoelectronic integration of a GaAlAs laser, a field-effect transistor, and a photodiode" Appl Phys Lett, Vol 44, No 10, 1984, (941-943)

94 T Sanada, S Yamakoshi, O Wada, T Fujii, T Sakurai and M Sasaki, "Monolithic integration of an AlGaAs/GaAs multiquantum well laser and GaAs metal-semiconductor field-effect transistors on a semi-insulating GaAs substrate by molecular beam epitaxy", Appl Phys Lett, Vol 44, No 3, 1984, (325-327)

95 H Matsueda, S Sasaki, and M Nakamura, "GaAs optoelectronic integrated light sources", J Lightwave Tech, Vol LT-1, No 1, 1983, (261-269)

96 J Carney, M Helix, R Kolbas, S Jamison and S Ray, "Integrated optoelectronic transmitter" Proc SPIE, Vol 208, (121-127)

97 S Miura, O Wada, H Hamaguchi, M Ito, M Makiuchi, K Nakai and T Sakurai,"A monolithically integrated AlGaAs/GaAs pin/FET photoreceiver by MOCVD" IEEE Electron Dev Lett, Vol EDL-4, No

10. 1983, (375-376)

98 R F Leheny, R E Nahory, M A Pollack, A A Ballman, E D Beebe, J C DeWinter and R J Martin, "Integrated InGaAs pin/fet photoreceiver", Electron Lett, Vol 16, 1980, (353-355)

99 R M Kolbas, J Abrokwah, J K Carney, D H Bradshaw, B R Elmer and J R Biard, "Planar monolithic integration of a photodiode and a GaAs preamplifier", Appl Phys Lett, Vol 43, No 9, 1983, (821-823)

100 I Hayashi, " 'OEIC': Its concepts and prospects", Proc 4th Int Conf on Integrated Optics and Optical Fibre Communication, Tokyo, 1983, paper 28B2-4

101 B K Nayar, D R Smith, C S Yoon and J N Sherwood, "Growth and assessment of highly nonlinear organic optical materials" Proc Conf on Lasers and Electro-optics, Anaheim, 1984, paper FO1

DISCUSSION

H.M.Gibbs, US

You mentioned one db per centimetre of loss in GaAs at 1.56 microns. Could you tell us what the origin of the free-carrier absorption is and how it changes as you move the wavelength toward the band edge?

Author's Reply

Clearly losses get higher as you go near the band edge. In our devices we try and work as far from the band edge as possible. It's clear also that losses increase as you go to much longer wavelengths because of increased free-carrier absorption. We want to operate at an optimum position for low loss, and we haven't studied the effects of getting higher loss by going nearer the band edge.

MULTIPORT OPTICAL DETECTORS

N. G. Walker
J. E. Carroll

Cambridge University Engineering Department,
Trumpington Street, Cambridge,
England, CB2 1PZ.

ABSTRACT

At microwave frequencies the application of multiport junctions to making phase and amplitude measurements is well established. In this technique unknown signals are combined in a multiport to give a number of output signals. Monitoring the power at each output port then enables the required signal information to be deduced. Here we investigate the equivalent optical measurements made when a number of light beams are mixed in a lossless junction and then detected using photoemissive devices. We call this optical multiport detection by analogy with the microwave theory. Applications of some optical multiports to various phase and amplitude measurement problems are discussed. The quantum theory of multiport detection is introduced, and the quantum operator corresponding to the complex amplitude measurement made by a particular multiport is presented. It is not known how to realise practically a measurement of the quantum maximum likelihood phase estimator, however the performance of the multiport operator in measuring phase comes within seventy percent of this limit.

Introduction

Central to optical communication and optical computing systems is the combining and detection of light beams. In a heterodyne or homodyne communication system one mixes a local oscillator with the incoming signal before detection, using for example a beamsplitter or a fibre coupler. In a computer one may be interested in splitting an optical signal into many different channels, or in combining signals on a detector to realise a logic element.

We intend to examine a large class of lossless beam combining which we call optical multiport detection. The name is coined from microwave frequencies where the theory of multiports is well established. At these frequencies there has been recent interest in using various multiport junctions for relative phase and amplitude measurements [1,2] and these techniques are now being used to construct automatic network analysers [3]. It seems that some insight may be gained into the equivalent optical measurements by bearing the microwave ideas in mind. These microwave multiport measurement techniques are particularly suited to realisation at optical frequencies since they rely solely on power measurements from several ports, which transfers to photon counting in the optical case. One may thus investigate how photon number measurements relate to amplitude and phase measurements and what implications this has on, for example, the design of receivers in coherent optical communication systems.

We will firstly show how the main features of microwave multiport theory carry over to optical frequencies, and consider examples of homodyne circuits making complex amplitude measurements. We then briefly describe a multiport we are building to demonstrate these principles. The quantum theory of multiport measurements is introduced and the result of applying this to our experiment is presented. We finally compare the quantum optimum phase measurement on a coherent state with that achievable by our multiport.

Semiclassical theory

First we review some of the main features of microwave multiport theory as this will carry over in a very similar way to optical frequencies. Indeed, for most cases of current practical interest the only extension required at optical frequencies is the introduction of shot noise resulting from the detection of discrete photons. In the semiclassical model this noise is assumed to be a Poisson process, with the detection rate proportional to the incident power level, however, it will be pointed out later that this is not an accurate enough description of the detection statistics for all possible quantum states of the signal field. It does, however, give the right predictions for a certain class of states which conveniently include the outputs from most presently available light sources.

Consider a general waveguide junction with M ports as shown in figure 1. We assume for simplicity that only one propagation mode in each guide is supported and that we

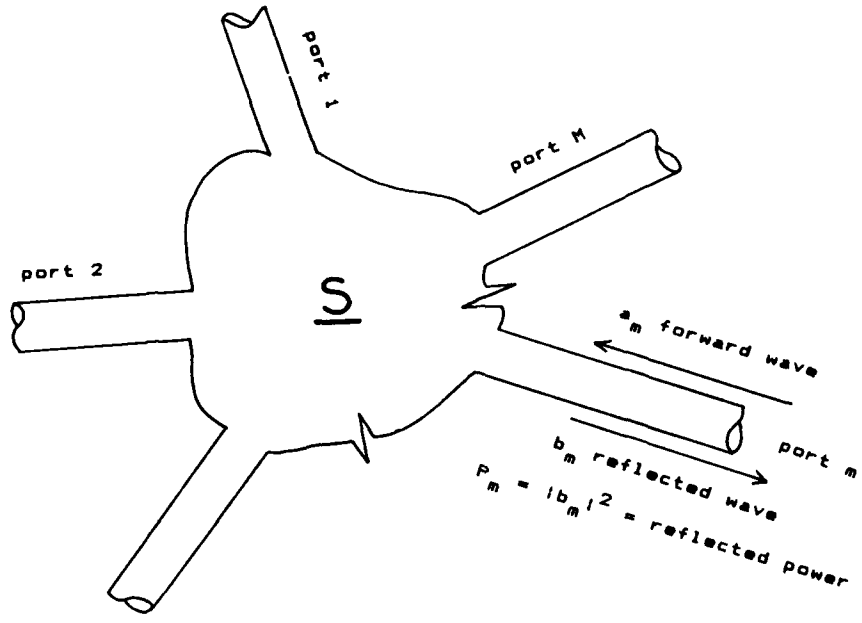


Figure 1. A general microwave M-port junction. The reflected waves are related to the input waves by the scattering matrix; $\underline{b} = \underline{S}\,\underline{a}$. If the multiport is lossless then the scattering matrix is unitary.

need only consider a single frequency signal. The complex amplitude of the input to port m is denoted by the mth component of the vector $\underline{a}$ and similarly the reflection from port m has a complex amplitude represented by the mth component of $\underline{b}$. The output complex amplitudes are related to the input complex amplitudes by the scattering matrix $\underline{S}$. If the multiport junction is lossless power is conserved and the scattering matrix is unitary. We have then

$$\underline{b} = \underline{S}\,\underline{a} \qquad \text{with} \qquad \underline{S}^{\dagger}\,\underline{S} = \underline{S}\,\underline{S}^{\dagger} = 1 \tag{1}$$

The power output to a detector at the mth port is

$$P_m = |\underline{b}_m|^2 = |(\underline{S}\,\underline{a})_m|^2 \tag{2}$$

An important example of a four port circuit is shown in figure 2a. This is the directional coupler which forms the main building block in a lot of microwave circuits. A suitable scattering matrix for the directional coupler is

$$\underline{S} = \begin{bmatrix} 0 & 0 & p & -q \\ 0 & 0 & q & p \\ p & q & 0 & 0 \\ -q & p & 0 & 0 \end{bmatrix} \qquad \text{with} \qquad \begin{matrix} |p|^2 + |q|^2 = 1 \\ pq^* - qp^* = 0 \end{matrix} \tag{3}$$

Figure 2b shows how a cube beamsplitter behaves as an optical directional coupler. Directional couplers may also be made in integrated form or by bringing two optical fibres into close contact. In fact it can be shown that any completely matched lossless reciprocal four port behaves as a directional coupler [4], so it is not surprising that they can be realised in many different ways.

A typical optical heterodyne or homodyne detection scheme uses a beamsplitter (directional coupler) to mix the local oscillator and signal beams on the detector surface. This is shown schematically in figure 3. We assume equal power beamsplitting so that

$$|p|^2 = |q|^2 = 1/2 \tag{4}$$

The signal and local oscillator combine to give outputs b_c and b_d. The detectors then

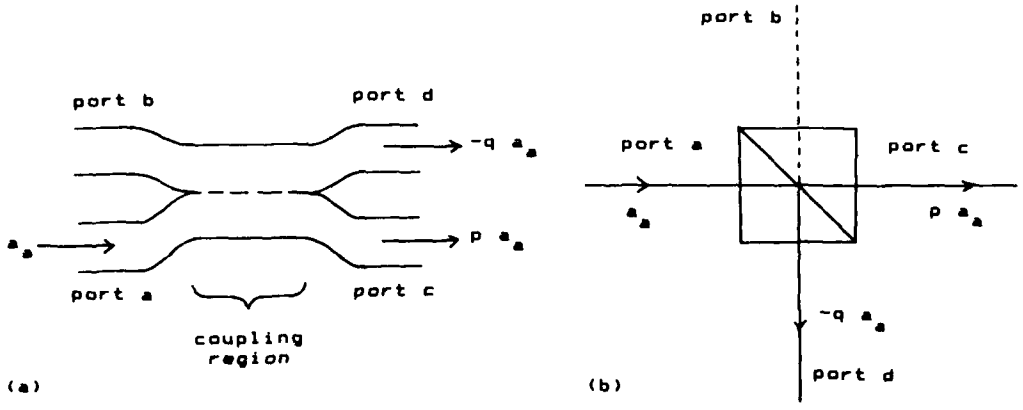


Figure 2. The directional coupler is an important building block in microwave circuits. (a) shows schematically how it may be realised at microwave frequencies by coupling the fields through small holes between guides. If the device is lossless then $|p|^2+|q|^2 = 1$. (b) shows how a cube beamsplitter behaves as a directional coupler at optical frequencies.

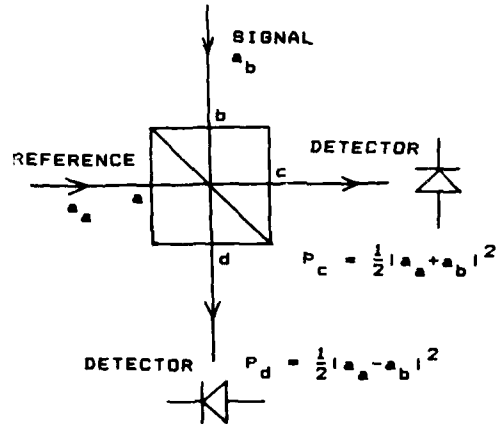


Figure 3. A single beamsplitter is used to mix the signal and local oscillator in a homodyne or heterodyne communication system. This is a balanced configuration where the signal and local oscillator are split equally between both detectors. We have then $P_c-P_d = 2|a_a||a_b|\cos\theta$ where θ is the phase of the signal relative to some zero.

measure the powers P_c and P_d.

Using equations 3 and 4 we have

$$P_c = |b_c|^2 = \tfrac{1}{2}\left(|a_a|^2 + |a_b|^2 + 2|a_a||a_b|\cos\theta \right) \qquad (5)$$

and

$$P_d = |b_d|^2 = \tfrac{1}{2}\left(|a_a|^2 + |a_b|^2 - 2|a_a||a_b|\cos\theta \right) \qquad (6)$$

so that

$$P_c - P_d = 2|a_a||a_b|\cos\theta \tag{7}$$

Where θ is the phase difference of the two inputs relative to a zero determined by the beamsplitter.

If we monitor the local oscillator power so that $|a_a|$ is known then we can deduce the 'in phase' component of the signal complex amplitude. In a microwave system an extra sampling port is often included in the circuit to do the monitoring, however, if the local oscillator is stable enough then it may be omitted.

In many applications it suffices to measure only one component of the signal complex amplitude relative to the local oscillator. This is the case in a two level phase shift keying system, or a homodyne amplitude shift keying system, provided that the phase of the local oscillator is locked to the zero phase of the signal. In other applications, however, we may need to measure both components of the complex amplitude. This is the case if the local oscillator phase is not known, or if we have a multilevel phase shift keying system.

If we wish to measure both 'in phase' and 'quadrature' components of the complex amplitude then we have to go to a more complicated multiport arrangement. The minimum number of power measurement ports required to do this is three, and a possibility for realising this is shown in figure 4a. However, an alternative configuration using four power measurements is shown in figure 4b. This is an eight port circuit, built from beamsplitters, in which four ports are input and four ports are output. This particular circuit can be simply understood as two of the single beamsplitter schemes arranged so that the relative phase of the local oscillator and signal is shifted by $\pi/2$ in each case. One combination of signal and local oscillator then measures the and the other combination measures the 'quadrature' component.

The $\underline{S}$ matrix for the eight port is assumed to be of the form

$$\begin{bmatrix} b_a \\ b_b \\ b_c \\ \cdot \\ \cdot \\ \cdot \\ b_h \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & 0 & \multicolumn{4}{c}{\underline{S}_{backwards}} \\ 0 & 0 & 0 & 0 & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & 0 & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & 0 & \cdot & \cdot & \cdot & \cdot \\ \multicolumn{4}{c}{\underline{S}_{forwards}} & 0 & 0 & 0 & 0 \\ \cdot & \cdot & \cdot & \cdot & 0 & 0 & 0 & 0 \\ \cdot & \cdot & \cdot & \cdot & 0 & 0 & 0 & 0 \\ \cdot & \cdot & \cdot & \cdot & 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} a_a \\ a_b \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ a_h \end{bmatrix} \tag{8}$$

Reciprocity will imply relations between $\underline{S}_{forwards}$ and $\underline{S}_{backwards}$. If only ports a to d are excited then we need only consider $\underline{S}_{forwards}$ which we will henceforth denote simply by $\underline{S}$. By suitable choice of the optical path lengths between the beamsplitters the forwards scattering matrix may be chosen as

$$\underline{S} = \frac{1}{2}\begin{bmatrix} 1 & 1 & -j & -j \\ 1 & -1 & -j & j \\ 1 & -j & j & 1 \\ 1 & j & j & -1 \end{bmatrix} \tag{9}$$

If the local oscillator and signal have complex amplitudes a_a and a_b respectively and there is no input into ports c and d then the differences $P_e - P_f$ and $P_g - P_h$ give the in phase and quadrature components of the signal respectively.

$$P_e - P_f = |a_a||a_b|\cos\theta \tag{10}$$

$$P_g - P_h = |a_a||a_b|\sin\theta \tag{11}$$

This multiport can therefore be used to determine both amplitude and phase, so we have made a homodyne measurement of the complex amplitude of the signal relative to the local oscillator.

Measuring both quadratures of the signal has some points which may be useful in certain communication problems. The four power measurement scheme provides the sensitivity of a homodyne system regardless of the phase of the incoming signal. In a conventional homodyne system the phase of the local oscillator must track that of the signal exactly in order to maintain the sensitivity, however, in this example either the sine term or the cosine term will be sensitive to the presence of a signal regardless of its incoming phase. In fact if the two terms are squared and added

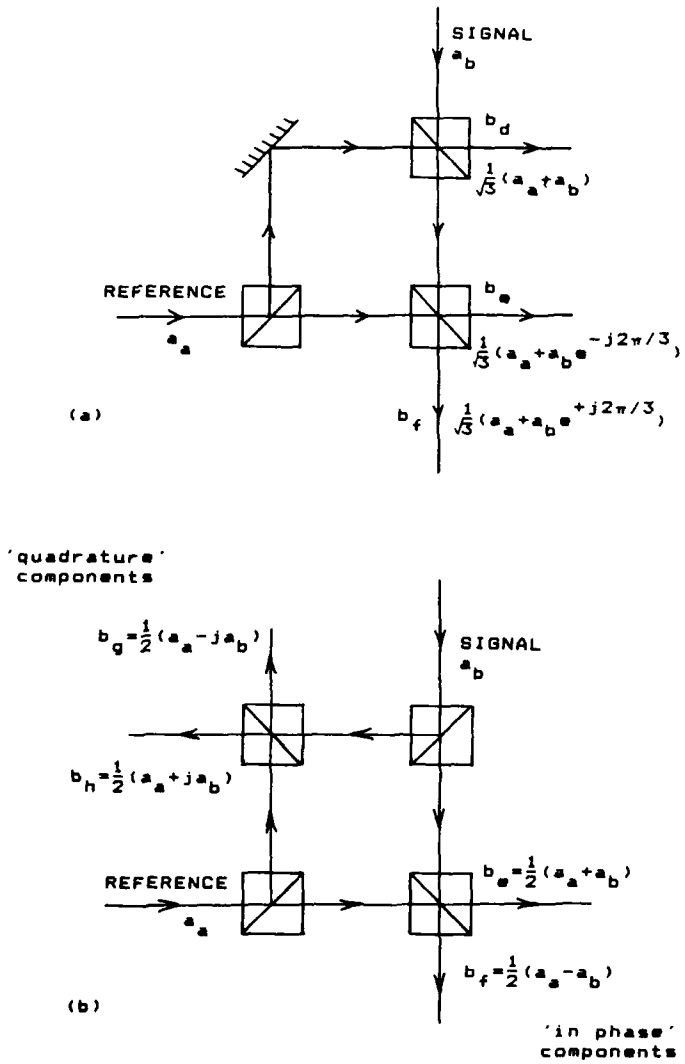


Figure 4. If both `in phase' and `quadrature' components of the signal are required then we must make at least three power measurements (in a lossless multiport). (a) shows a configuration built from beamsplitters which uses the minimum number of detectors. This example uses two equal power splitters and one two thirds - one third splitter. (b) shows a simpler configuration using four detectors where one half of the circuit makes the `in phase' measurement and the other half makes the `quadrature' measuement.

then we obtain immediately the amplitude squared of the signal.

There is present interest in phase shift keying communications systems using homodyne detection, however, there is great practical difficulty in stabilising the lasing frequency of an independent local oscillator and then locking its phase to that of the signal zero. The four power scheme could be used to detect phase shifts in the incoming signal without the necessity for the local oscillator phase to be locked to the signal, provided the rate of phase change due to modulation was much greater than that due to the difference frequency between the signal and local oscillator.

There is also a requirement for making accurate phase measurements in optical fibre gyroscopes, and 3 x 3 directional couplers have already been used to improve their

sensitivity [5]. This is another application of multiport measurement principles.

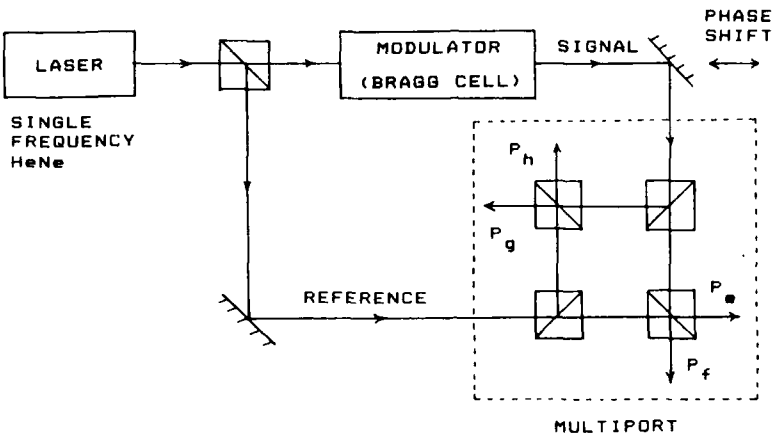


Figure 5. Experimental arrangement to demonstrate multiport measurement principles. This is a homodyne system with the local oscillator and signal beams taken from the same light source.

Experimental Work

In order to investigate a simple multiport circuit experimentally we have chosen the four power measurement scheme discussed here and are building a homodyne system to make relative amplitude and phase measurements. The arrangement is shown in figure 5. To avoid the problems of stabilising the frequencies of the local oscillator and signal we have derived both from the same source, a one milliwatt single frequency helium-neon laser. We have used bulk optics and free space propagation in our initial experiments for flexibility and to avoid the problems of maintaining polarisation, although it is envisaged that a production detector would be made in integrated form. After splitting the light into local oscillator and signal beams the signal is attenuated using neutral density filters and modulated using a Bragg cell. The two are then recombined in the block of four beamsplitters comprising the eight port detector. Photodiodes detect the four output powers. Since the wavelength of light is 0.6 microns and we intend to measure phase to an accuracy of a few degrees the system must be stable to about about 10 nanometers. To minimise these difficulties we have chosen a balanced system where only the relative path lengths are important. Piezo electric crystals are used to maintain the position of the beamsplitters to fine resolution, and these will eventually be controlled by a microcomputer which will keep the system in alignment and calibration. A Piezo electric movement is also used to vary the phase of the signal beam. The Bragg Cell limits modulation bandwidth to 3 MHz, so the first experiments have been performed with 1 microsecond pulses at about 15 KHz repetition rate. The statistics of these measurements will eventually be fed into the microcomputer to be analysed. To simulate a viable communication system work would have to be done at higher frequencies.

Initial trials show accuracy to about five degrees of phase and ten percent in amplitude with a signal energy corresponding to about 10^5 photons in each pulse, although the system has not yet been properly calibrated. At the time of writing no work has been done on reducing the noise.

Quantum Theory

Inevitably we are interested in the noise performance of multiports and the parameters dictating this. At optical frequencies the fundamental noise source is the quantum mechanical statistics of the detection of photons, so a few results of the the quantum theory are presented below, although a detailed analysis would be out of place here.

The quantum mechanical noise in making a measurement is determined by both the state of the system and the operator corresponding to that measurement. In a communication

channel we imagine the system we are measuring to be the electromagnetic field entering the detector, the state of which depends on the message sent by the transmitter. We consider the quantum measurement to be complete when the signal has been amplified in some way, for example by photoemissive detection and avalanche gain in a photomultiplier tube, so that it can be considered classically, and in principle measured as accurately as we wish.

There exists a well developed theory of quantum optimum detection which is concerned with choosing the best measurement operators, given a set of possible system states with their probability distribution, and a cost function for the noise of the measurement. An excellent coverage of most of the results of this theory is given by Helstrom [6] (see also [7] for a shorter introduction). Unfortunately not many results exist for the translation of these abstract measurements into a physically realisable experiment, however, some important cases have been solved [8,9,10,11, also 6 pp. 160-165]. Even for measurements for which the physical realisation of the optimum operator is not known, the abstract theory still provides a useful yardstick with which to evaluate the performance of real systems making the same measurement.

We would like to know what quantum operator corresponds to a given multiport measurement. With this operator we can evaluate the multiport noise performance in abstract quantum terms, gain a better understanding of the sources of noise, and compare the performance of this measurement technique with that of the absolute quantum optimum.

As a starting point for our calculations we take an important theorem provided by Yuen and Shapiro [11], which gives a stepping stone between the abstract quantum measurements and the quantities we observe with an electronic detection system. This theorem identifies the photoelectron counting distributions from photoemissive detectors, such as photodiodes or photomultiplier tubes, with the abstract quantum measurement of the photon number operator. The equivalence of these two measurements has been assumed intuitively for a long time, but Yuen and Shapiro with work based on that of Kelly and Kleiner [11] have given a proof [13,11].

We consider a detector array of unit quantum efficiency with the mth surface detecting the photons from the mth port of our multiport. The photon number operator for the output field in this port is then

$$\hat{N}_m = \hat{b}_m^\dagger \hat{b}_m \tag{12}$$

where $\hat{b}_m$ is the photon annihilation operator for the single mode output field in the mth port (proportional to the complex electric field operator) and $\hat{b}_m^\dagger$ is its Hermitian conjugate, or the creation operator.

In the quantum multiport problem we have a set of input fields with quantum states described by their density operator, and a set of detectors which operate on combinations of these fields. We assume that the electric field operator transforms between the inputs and outputs of the multiport in the same way that the classical fields do: via the scattering matrix. Since the photon annihilation operators are related to the electric field operators by a normalisation constant then these also transform via the scattering matrix. If $\underline{\hat{b}}$ are the set of photon annihilation operators for the output fields and $\underline{\hat{a}}$ are the set of annihilation operators for the input fields, with Hermitian conjugates $\underline{\hat{a}}^\dagger$ and $\underline{\hat{b}}^\dagger$ respectively, then we have

$$\underline{\hat{b}} = \underline{S}\ \underline{\hat{a}} \quad \text{and} \quad \underline{\hat{b}}^\dagger = \underline{S}^*\ \underline{\hat{a}}^\dagger \tag{13}$$

The justification for this is :-

- The quantum field operators transform in exactly the same way as the classical multiport fields. This is analogous to the Heisenberg equations of motion for operator transformations in time.
- The behaviour of the quantum multiport and classical multiport are exactly similar when the input fields are a random superposition of coherent states. The coherent state is the closest quantum analog to the classical stable oscillation.
- The field commutator between the output fields is the same as between the input fields. That is the field operators for different output ports are independent (they commute). Von Neumann states that any canonical transformation (one which preserves the commutator) can be expressed as a unitary transformation.
- This assumption is compatible with those already made in the literature about the nature of beamsplitting.

The photon number operator for each port is then

$$\hat{N}_m = (\underline{S}\ \hat{a})_m^\dagger\ (\underline{S}\ \hat{a})_m \tag{14}$$

6. FAST MODULATION OF SEMICONDUCTOR LASER DIODES

The high dynamic range of the picosecond photoconductive switches allows the use of ultrashort electrical output pulses for a direct modulation of semiconductor lasers. A GaAs:Cr-picosecond switch has been used to directly drive a GaAs:GaAlAs buried heterostructure laser [42] to study the transient infrared laser emission under extremely fast carrier injection (see figure 7).

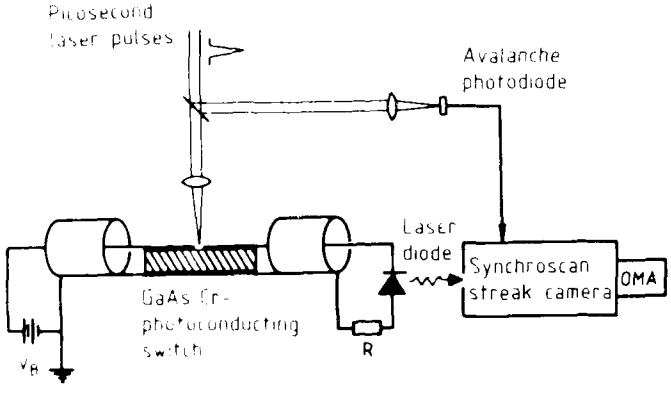


Figure 7
Arrangement for direct gain switching of a laser diode by the picosecond pulse train of a GaAs:Cr photoconductive switch.

The laser diode under rapid gain modulation is emitting infrared pulses comparable in width and strongly synchronized to the electrical driving pulses. The short output pulses of the semiconductor laser have been analysed in time by a Synchroscan streak camera with a temporal resolution < 10 ps.
As has been elucidated by a simultanuous spectral analysis the emission of the buried heterostructure laser under ultrafast modulation is always multimode in contrast to cw driving conditions where clean single longitudinal mode operation is observed (figure 8). Only by applying simultanuously a DC bias the laser emission remains single mode. However, the emitted optical pulses are considerably broadened in time /44/. This temporal pulse broadening can be attributed to diffusion damping: homogeneously injected carriers as a consequence of spatial hole burning diffuse preferably into the center part of the transverse mode volume and give a contribution to stimulated emission even after switching off of the electrical driving pulse. This pulse broadening effect which has also been observed in gain guided lasers /48/ is of importance for the development of high-speed optical communication systems, where transmission rates of > 1 Gbit/s are planned in the near future. The difficulty of pulse broadening possibly can be overcome by injection locking /49/ or by the use of ultrashort laser diode cavities /48/.

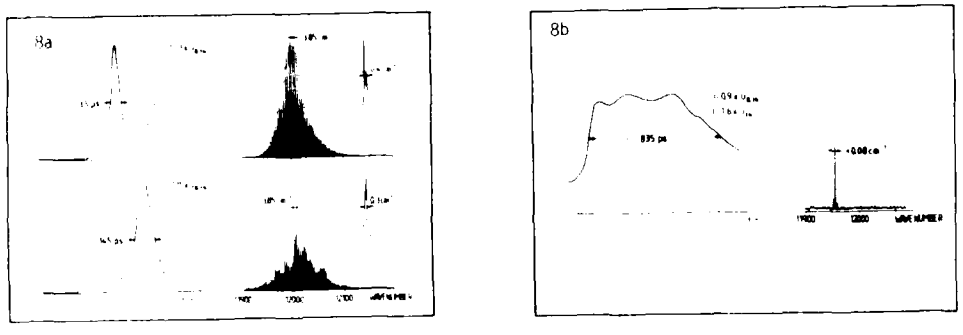


Fig.8: Temporal and spectral output for unbiased (8a) and dc-biased GaAs/GaAlAs BH-laser (8b). For the unbiased operation (8a) the driving pulse amplitudes are 1.15 and 1.53 times the threshold value $U_{B,th}$. For the the dc-biased operation (8b) the dc current is 1.6 times the threshold current I_{th}, superposed by pulse amplitudes of 0.9 times the threshold value $U_{B,th}$.

The combination of laser-activated optoelectronic switch with a Synchroscan streak camera (see figure 7) provides a very useful tool to study picosecond transients in light emitting diodes and semiconductor lasers at GHz modulation rates.
The electrical driving pulses can be manipulated optically to form high frequency electrical-bit-patterns (see next chapter) suitable for driving semiconductor lasers at high-bit rates (> Gbit/s) under suppression of the so-called pattern effect /50/ and under control of relaxation oscillations.

5. HIGH SPEED DEVICE CHARACTERIZATION

The arrangement depicted in figure 4 can be extended to a general optoelectronic sampling system for the characterization of high speed electronic devices:
On one hand the picosecond photoconductive switch can be considered as source for ultrashort electrical pulses, on the other hand it can be used as ultrafast sampling gate for the measurement of extremely short electrical signals. Both capabilities can be combined if two photoconducting devices are connected in series with a high speed device to be investigated in between as shown in figure 6.

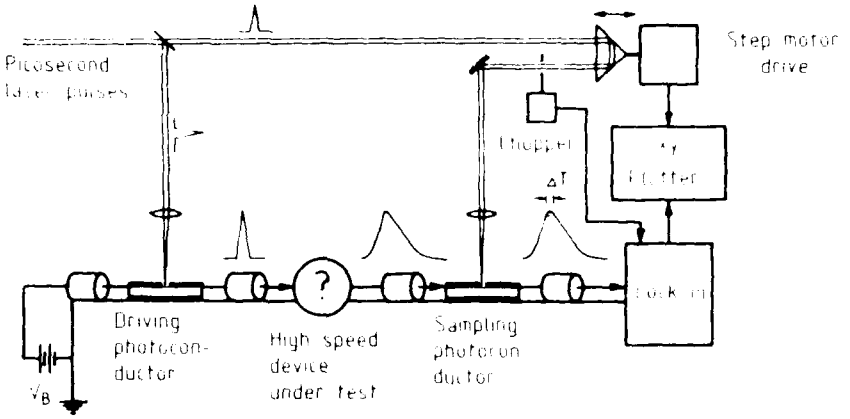


Figure 6: Optoelectronic sampling system for high speed device characterisation

The high speed device under test is driven by the electrical output pulses of a fast photoconductor. The electrical response of the device is sampled by a second photoconductor which is strongly synchronised to the driving photoconductor since both are activated optically by the same picosecond pulse train. The measured crosscorrelation curve Q(T) is a convolution of the device response, the sampling photoconductor response and the optical pulse profile. The time-resolution of this optoelectronic sampling system is by an order of magnitude higher than that of conventional electronic sampling systems. With the arrangement sketched in figure 6 for the first time the impulse response (25 ps FWHM) and the internal delay (19 ps) of a high speed FET has been directly measured /27/. The most widely used conventional approach to measure the speed of fast FETs is to connect a number of them in a tandem configuration in a ring oscillator to a string of identical logic gates and inverters. From the response of the ring oscillator the individual device properties can be deduced only indirectly. The time resolution of picosecond optical electronics has been extended to the subpicosecond regime by an electrooptic sampling technique which uses a travelling wave Pockels cell in conjunction with a subpicosecond laser system /43/.
The performance of the picosecond optoelectronic sampling is compared in table 3 to that of a conventional pure electronic measurement scheme and to that of a more future approach of "superconducting electronics" instrumentation.
The most attractive features of picosecond optoelectronics as high speed, large dynamic range, negligible jitter and very high sampling rates allow an efficient signal averaging for the measurement of extremely low level signals without requiring high speed external electrical circuitry. Unlike to a superconductive electronics approach, which uses Josephson junctions and SQUID's (Superconducting Quantum Interference Devices) as ultrafast gates, the picosecond optoelectronic measurements are not restricted to extremely low operating temperatures.

Table 3
High speed electronic measurements schemes (after /25/)

Method	Conventional Electronics	Superconducting Electronics	Picosecond Optical Electronics	
Instrumentation	Pulse Generator + Sampling Scope	Liquid-helium environment Josephson-junctions SQUID logic gates	cw - or single pulse (s) picosecond laser ultrafast optoelectronic switches	
Gating time/speed	≈25 ps	≈20 ps	cw ≈1 ps	s ≈40 ps
Dynamic range	10^{-3}-10 V	10^{-5}-10^{-1} V	10^{-5}-10 V	10-10^{4} V
Sampling Rates	≈10^{6} Hz	≈10^{8} Hz	≈10^{10} Hz	bias voltage dependent

4. OPTOELECTRONIC SAMPLING TECHNIQUE FOR HIGH SPEED MATERIAL CHARACTERIZATION

The simplest way to determine the transient response of a photoconducting material is the analyis of the electrical signal monitored at a fast sampling oscilloscope, which is a convolution of the optical pulse width, the impulse response of the sampling head, connectors, transmission lines and the time material response. Since most of these are not known precisely, it is difficult to estimate the true detector response by deconvolution. For characterizing materials and devices with response times below 25 ps conventional electronic instrumentations cannot be used.

A new optoelectronic sampling technique originally developed by D. Auston and coworkers /4,17/ which uses ultrafast photoconducting devices, however, is capable to measure electronic signals with picosecond time resolution.

The principle of this measurement technique for material characterization is sketched in figure 4. Two photoconductors are connected in series on the same top electrode of a microstrip transmission line. A pulse train of a mode-locked picosecond laser is split up in two parts to activate the first gap photoconductor and after a variable delay T the second gap photoconductor on the microstrip transmission line top electrode. The electrical signal of the first photoconductor propagates along the microstrip electrode and can be sampled quasi-optically at the second gap by the delayed picosecond pulse with a gating-time comparable to the optical pulse width (typically 3-5 ps in synchronously pumped mode-locked dye lasers). The integrated charge Q(T) sampled at the output side of the device as a function of time delay T is proportional to the convolution of the responses of the two photoconductors /4/

$$Q(T) \propto \int U(t)U(t+T)dt \qquad (4)$$

If the two photoconductors are identical and the two gaps are of the same dimensions Q(T) represents the autocorrelation function of the material response. From the half width (FWHM) of the autocorrelation function T_{ACF} the real pulse width T_P of the electrical signal can be deduced by

$$T_P = A \cdot T_{ACF} \qquad (5)$$

where A is the factor in the order of unity which depends from the pulse shape ($A = 1/\sqrt{2} = 0.7$ for a Gaussian pulse profile.

Figure 5 shows the autocorrelation functions measured on SOS-photoconductors irradiated with different doses of 150 KeV O^+-ions /11,46/.

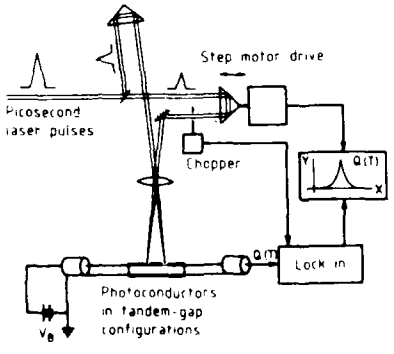


Figure 4
Optoelectronic sampling technique for high speed material characterization /4,17/

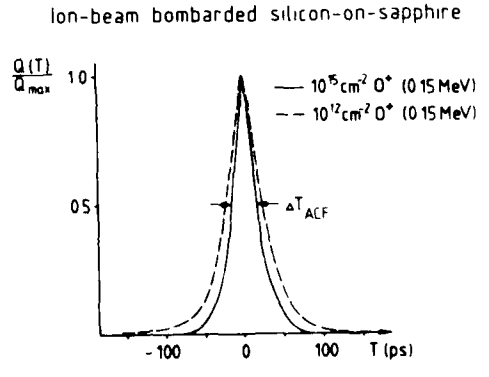


Figure 5
Autocorrelation traces obtained by optoelectronic sampling of O^+-ion beam bombarded silicon-on-sapphire /46/

In practice complications arise from electrical reflections by impedance-mismatch at the coax-microstrip connections.

Additionally it has to be taken into account that unsymmetrical electrical pulses propagating along microstrip transmission lines can be broadened or compressed as a consequence of frequency dispersion /47/.

By an integrative lock-in measurement of the transmitted charge Q(T) with typical integration times of 1s a signal-to-noise ratio of 10^4 has been achieved for mV-electrical signals /1,17/.

Table 2
Comparison of performance parameters of standard semiconductor junction photodiodes and new photoconductive photodetectors.

Material	Sensing type	Rise-time (ps)	Response-time (ps)	Spectral range (μm)	References
Si	PIN	35	50	0.3-1.1	
Si	APD	85	120	0.3-1.1	
GaAs/GaAlAs	APD	75	100	0.4-0.9	
Ge	PIN	50	75	0.5-1.8	
a-Si	PC	<5	<10	0.3-1.1	3,4
Si (RD)	PC	<5	<5	(<0.01)-1.1	17,40
GaAs:Cr	PC	<5	60	0.4-1.0	8 - 12
InP(RD)	PC	<5	<5	0.6-1.5	18,19
p-InGaAs	PC	<40	70	1.0-1.7	14
a-Ge	PC		50		7
CdSe	PC		20		22

Abbreviations:
PIN...p-i-n-type photodiode
APD...Avalanche-type photodiode
PC ...Photoconductive photodetector
RD ...Radiation damaged material

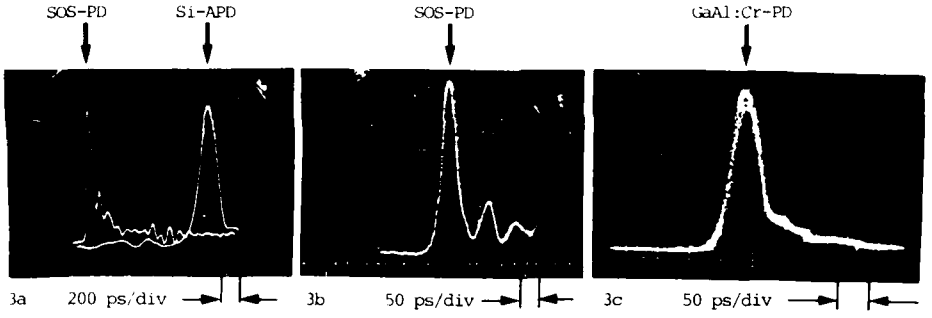


Figure 3
Photoresponse of picosecond photoconductive detectors activated by the pulse-train of a synchronously pumped mode-locked dye laser:
a) Comparison of photoresponse of SOS-photoconductive detector with that of a commercially available Si-APD; b) Response of SOS-photoconductive detector; c) Response of GaAs:Cr photoconductive detector.

As can be seen from table 2 the advantage of the new photoconductive detectors over conventional junction type photodetectors is their high speed performance.
Figure 3a) shows the photoresponse of a photoconductive SOS detector (left signal) as compared to that of a commercially available "fast" Si-avalanche-photodiode when a picosecond pulse train of a synchronously pumped mode-locked dye laser (individual pulse width = 4 ps FWHM) is monitored on a fast sampling oscilloscope (Tektronix 7904/7T11). As can be seen from fig. 3b) where the response of the SOS-photodetector is monitored at an expanded time scale (50 ps/div), the signal is limited by the specified risetime of the fast sampling head (25 ps). The actual pulsewidth of the photoconductive detector response has been determined by an optoelectronic autocorrelation technique (see next chapter) to be <15 ps.
The undulations observed at the trailing edge of the signal are due to electrical reflections from the microstrip-to-coax connections.
Figure 3c) shows for comparison the temporal response of a GaAs-Cr photodetector. Again the observed risetime is limited by the speed of the sampling head and the cable connecting network. The pulsewidth is determined to be about 60 ps (FWHM). The trailing edge of the electrical pulse shows an undesired tail which can be attributed to multiple trapping of the photocreated carriers.

Photoconducting materials and their applications for picosecond optoelectronic switching

Table 1

Material	Technology	Switching Speed* [ps]	Switching Amplitude* [V]	Applications	References
Si(intrinsic)		RT<10	10^2 (s)	High voltage switching	1, 37
Si:Au		RT<40	10^4 (s)	Microwave generation	2, 32
				Synchronisation of streak camera	29, 38
				Jitter-free driving of Pockels cell	30, 42
a-Si	EBE	<10	10^{-2} (cw)	Material characterization	1 - 6
	CVD	20			5
	GD	200			11
Si(SOS)	RD(O^+,D^+,p^+)	<20	$5x10^{-1}$ (cw)	Picosecond photodetection	17,11,45
				Picosecond UV-x-ray photodetection	39,40
				Picosecond sampling	17,28
				High speed device characterization	27,28
				High-bit-rate electrical code generation	41
GaAs:Cr		RT<40	$8x10^3$ (s)	High voltage switching	10
				Microwave generation	31
			10^3 (cw)	Active mode-locking of dye laser	42
		RT<1	10^{-1} (cw)	Subpicosecond sampling	43
		60	2 (cw)	Direct modulation of laser diodes	12,44
			2 (cw)	Synchronous mode-locking of laser diodes	45
p-InGaAs		70	$5x10^{-2}$ (cw)	Material characterization	14
InP:Fe		70	$5x10^{-2}$ (cw)		13,20
		<100	10^3 (s)	High voltage switching	15
InP	RD(P^+)	<100		Infrared ps-photodetection	18
	RD($3He^+$)	RT<1		Picosecond photodetection	19
a-Ge	LA	50	$5x10^{-2}$ (cw)	Material characterization	7
CdS		<200	$5x10^1$		21
CdSe		20	$2x10^{-1}$ (cw)	Picosecond photodetection	22

Abbreviations:
Technology: EBE...electron-beam evaporation
CVD...chemical-vapor deposition
GD ...glow discharge
RD ...radiation damage (ion bombardement)
LA ...recristallisation by laser annealing

Switching speed: RT...rise time
Switching amplitude: (s)...single pulse operation
(cw)...high repetition rate operation

* Approximative values as given in the cited references or estimated values if not given explicitely.

3. PHOTOCONDUCTORS FOR PICOSECOND PHOTODETECTION

The most direct application of the high-speed photoconducting materials is their use as detector for picosecond light pulses.
The detection of such ultrashort pulses, however, is usually more difficult than their generation. In colliding-pulse mode-locking ring dye lasers light pulses with 60 femtoseconds duration ($6x10^{-14}$s) can be generated which can be measured only by nonlinear optical autocorrelation techniques. The fastest semiconductor junction detectors have response times which are by three orders of magnitude higher than the achievable time range of optical pulse techniques. Response times of 50 ps (FWHM) have been reported for Si-PIN photodiodes and ~100 ps (FWHM) for GaAs heterostructure avalanche photodiodes. The response times of the new type of photoconductive detectors are by one order of magnitude shorter, as can be seen by the comparison of performance parameters listed in table 2. The spectral response of the different photoconductors is ranging from X-ray (ion bombarded SOS films) /40/ to near infrared (p-InGaAs /14/).

In materials with a high intrinsic or externally induced density of recombination and trapping centers the electron-hole pairs which are generated by optical band-band absorption can recombine or can be trapped within some picoseconds (= 10^{-12}s) /33,35/. This kind of semiconductor materials in general is undesired in conventional microelectronics where high purity materials are required. The carrier lifetimes of pure direct gap or indirect gap semiconductors are typically ranging from nanoseconds to microseconds, the carrier mobilities can range up to μ = 1000 - 10000 $cm^2V^{-1}s^{-1}$ / 36 /.
Thin amorphous semiconductor films as a-Si generated by electron-beam evaporation (EBE) or glow discharge (GD) on insulating substrates generally have an intrinsic high density of recombination centers (10^{20} - $10^{16}cm^{-3}$) /5, 6 / and exhibit typical carrier lifetimes in the order of 1 - 100 ps / 1-6 /. Their carrier mobility is strongly reduced to values $\mu \leq 1$ $cm^2V^{-1}s^{-1}$ which leads to a poor photoresponse. Transport measurements of these a-Si films indicate that the transient mobilities are due to extended-state conduction / 5,6 /.
Better results can be obtained by radiation damaging of thin silicon-on- sapphire (SOS) layers, which exhibit similar high speed, but mobilities which are by one or two orders higher than those obtained in amorphous films. By a variation of the radiation damage dose (ion implantation or electron-beam bombardment) the carrier relaxation time can be tuned externally in a controlled way, while in amorphous films the transport properties are intrinsically given by the production process / 5,11/.
Figure 2 shows the variation of carrier lifetime τ(a) and carrier mobility μ(b) in silicon-on-sapphire (SOS) layers under different doses of O^+- or D^+-ion bombardment /11/.
By applying ion-particle beam implantation with decreasing particle energy on the same sample, the degree of damage can be dispersed uniformly throughout the entire film thickness (1 µm).

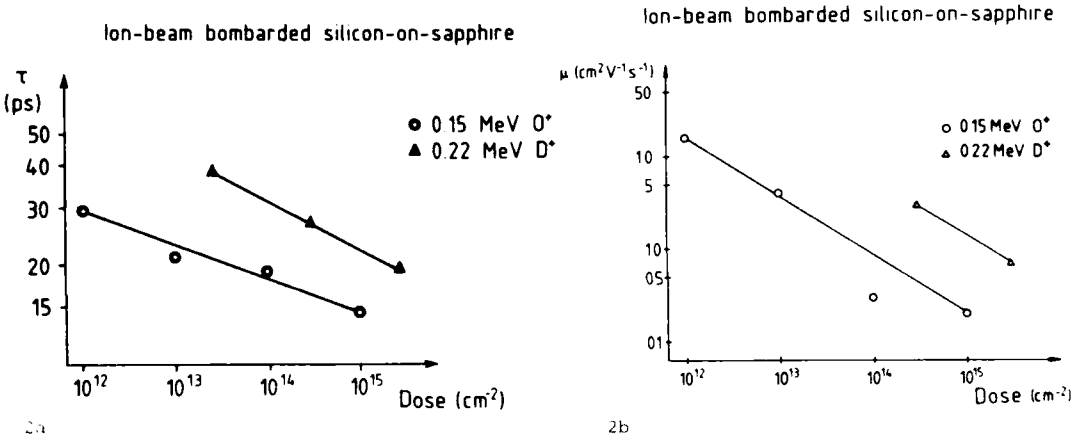


Figure 2
Variation of carrier relaxation time τ (fig. 2a) and carrier mobility μ (fig. 2b) versus doses of ion bombardment (150 KeV-O^+ ions and 220 KeV-D^+ ions respectively)

Significant higher mobilities are obtained in chromium(Cr)-doped GaAs (dopant concentration 10^{17} - 10^{18} cm^{-3}) which standardly is used as substrate material in some microelectronic devices. The photoexcited carriers are trapped within 60 - 70 ps due to the high Cr-concentration. The GaAs:Cr photoconductor can be driven at high repetition rates with a remarkable high dynamical range. Under activation with the continuous pulse train of a modelocked dye laser electrical pulses with a pulse width of ~ 60 ps (FWHM) and a signal amplitude of ~ 2 V can be generated at repetition rates of 100 MHz /11,12/. GaAs:Cr photoconducting switches have widely been used for high voltage applications: High voltage peaks up to 10 kV can be switched on with picosecond risetimes /10/. Similar good results have been obtained with Fe-doped or ion-implanted InP /13-19/.
Pioneering work on high speed photoconductive switching in silicon has been published by D. Auston and coworkers (Bell Laboratories, Murray Hill, USA) /1-6/. Important contributions to high voltage switching by photoconductive devices have been given by G. Mourou and coworkers (University of Rochester, N.Y., USA) /10, 26 /. Further work on picosecond photoconductivity and its applications has been published in a large number of papers during recent years.
Table 1 gives a general overview on photoconducting materials and their applications as reported by various groups.

2. PHOTOCONDUCTING MATERIALS FOR P COSECOND OPTOELECTRONIC SWITCHING

Starting points for the realisation of ultrafast optoelectronic switches on one hand have been the availability of picosecond laser pulses on the other hand the discovery and development of suitable photoconducting materials with extremely short carrier lifetimes. Figure 1 illustrates schematically the configuration and principle of operation of a simple photoconductive switching device:

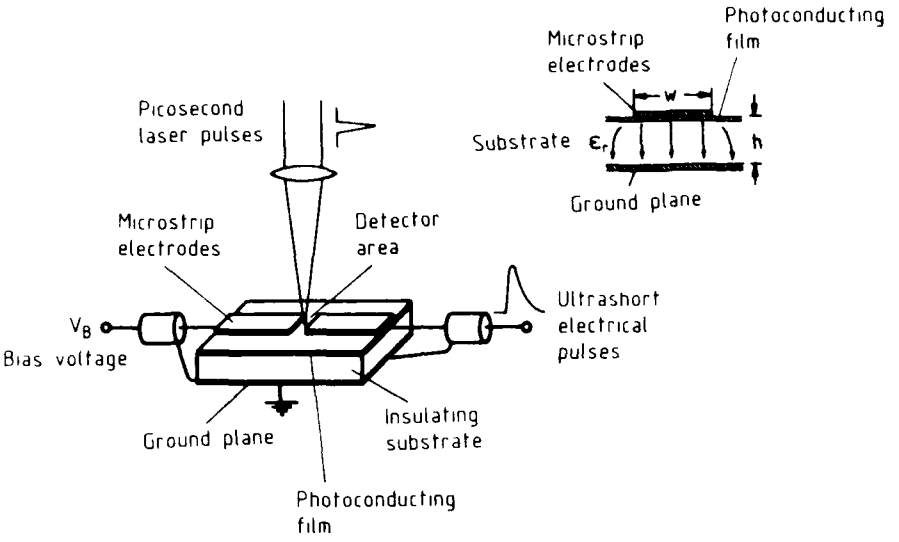


Figure 1: Principle of operation of high speed photoconductive device: A picosecond laser pulse is switching the photoconductivity in the microstrip electrode gap which leads to the generation of an ultrashort electrical pulse. The insert (right above) shows a cross section of the device (w...electrode width, h...substrate thickness, ϵ_r...relative permittivity)

It consists of a thin photoconducting film on an insulating substrate which is covered on both sides by metallic electrodes to form a microstrip transmission line configuration. The top side microstrip electrode exhibits a gap of appropriate dimensions (depending from operation voltage) which serves as detector area for the focussed picosecond laser pulses. For low voltage operation typical gap widths are ranging from 1 = 10 - 50 µm. The characteristic impedance of the microstrip transmission line is determined by the top electrode width w, the thickness of the substrate h and the relative permittivity of the substrate ϵ_r (see insert of figure 1) and is matched to the external 50 Ω circuitry. Both ends of the top electrodes are attached to standard microstrip-to coaxial connectors (e.g. 3 mm OSM) to provide an optimum high frequency transmission behaviour. At the input side of the photoconducting device a bias voltage V_B is applied. When the active detector area is hit by a picosecond light pulse the electrical conductivity in the gap is increased by several orders of magnitude and at the output side of the device appears an electrical pulse with an extremely short risetime. The peak amplitude and the temporal width are determined by the transient mobility and lifetime of the optically induced carriers. The impulse response of photoconductors in transmission lines is analysed in detail in ref. [1]. Qualitatively an infinitely short optical excitation pulse of energy E_P generates

$$N_O = \eta \cdot (1-R) \cdot [1-\exp(-\alpha \cdot d)] \cdot E_P/\hbar\omega \qquad (1)$$

electron-hole pairs, whereby η is the quantum efficiency for free carrier production, R the optical reflectivity of photoconductor, α the absorption constant, $\hbar\omega$ is the photon energy and d is the thickness of active photoconducting layer.
The induced photocurrent is given by

$$I_p(t) = 2 \cdot e \cdot \mu \cdot n_O \cdot E_B \cdot A \cdot \exp(-t/\tau) \qquad (2)$$

whereby τ is the carrier relaxation time (lifetime) and

$e = 4.8 \cdot 10^{-10}$ ESU		electron charge
μ	$(\mu_e + \mu_h)/2$	average carrier mobility (indices e and h designate electron and hole respectively)
$n_O = n_e = n_h = N_O/V$		carrier density (V...light absorbing volume)
$E_B = V_B/l$		electrical field strength on the top electrode gap (l...gap width, V_B...bias voltage)
$A = w \cdot d$		cross section of current leading region

According to relation (2) from the photo-response of the ultrafast device, measured as voltage signal

$$U(t) = U_O \cdot \exp(-t/\tau) = R_O \cdot I_p(t) \qquad (3)$$

on the input impedance R_O = 50 Ω of a fast sampling oscilloscope, can be deduced the transient mobility and carrier relaxation time of the photoconducting material.

PICOSECOND PHOTOCONDUCTIVE DEVICES FOR > 10 Gbit/s OPTOELECTRONIC SWITCHING*

G. VEITH, Standard Elektrik Lorenz AG, Research Center (ZT/FZNG), Hellmuth-Hirth-Str. 42, D-7000 Stuttgart 40, FRG

Semiconductor materials with a high density of recombination and trapping centers exhibit extremely short carrier lifetimes in the order of 1 - 100 ps and have been the base for the development of high speed optoelectronic switches. These devices are activated by picosecond laser pulses and can be driven nearly free of jitter with respect to the optical excitation pulses. They show some unique properties as picosecond risetimes and response times and can be operated within a relatively high dynamical range (10^{-5} - 10^{4} V). This article gives a review on the wide field of possible applications of the ultrafast photoconductive switches: They can be used as photodetectors for picosecond light pulses as well as sampling gates for the characterization of high speed electronic and optoelectronic devices. In some experiments which are discussed more in detail the author demonstrates the capability of this type of photoconductive switches for the generation of picosecond infrared pulse trains in laser diodes and for the generation of high-bit-rate electrical codes for use in Gbit/s optical communication and sensing systems, for logical switching and for testing purposes of high speed electronic instrumentation.

1. INTRODUCTION

The electrical conductance of a semiconductor material can be increased by several orders of magnitude if a picosecond light pulse is used to generate an electron-hole plasma of high density by optical band-band absorption. With modern mode-locked picosecond laser systems single pulse powers in the order of 1 kW (cw-systems) to 1 GW (single pulse systems) are available within the whole optical and near infrared spectrum so that extremely high plasma densities of 10^{18} - 10^{20} cm^3 can be achieved. D. Auston /1,2/ was the first to use high power picosecond laser pulses for the switching of photoconductivity in intrinsic silicon (Si). Due to the nanosecond lifetimes of electron-hole pairs in the intrinsic material the optically excited electrical pulse has been switched off by a second picosecond pulse at a longer wavelength (λ = 1.06 µm) which has been used to short-circuit the biased electrodes.

For the realisation of a high frequency operating photoconductive device it is, however, desirable to use a self limitating material, whereby the switching speed is defined by intrinsic properties of the semiconductor material itself. These properties of short carrier lifetimes and automatically switching off were found by different groups in several materials with a high density of recombination and trapping centers for the photoexcited carriers: amorphous materials (a-Si, a-Ge)/1 - 7/, heavily doped materials (e.g. GaAs:Cr, InP:Fe, Si:Au)/ 8-16/ and ion bombarded materials (Si-on-sapphire, InP) /17-20/. The new high speed materials in connection with the availability of picosecond- and femtosecond-laser systems /23,24/ led during recent years to a large number of papers in the new field of picosecond optical electronics /25,26/. This new technology uses the picosecond time resolution capability of ultrashort optical pulses for characterizing materials and devices which operate at such high speed that conventional electronic instrumentation is not capable to measure their responses /27,28/. As additional advantage the new photoconductive devices operate within a large dynamical range with negligible jitter allowing the measurement of low level signals (10 µV) as well as of high voltage peaks (10 kV) without significant loss of speed. This, furthermore, makes its use possible for a jitter free driving of electrooptic modulators or optoelectronic devices for a stable synchronisation of instrumentation in connection with time resolving optical picosecond experiments /26,29/. High voltage optoelectronic switches have also been used to synchronize Pockels cells for pulse compression of high power laser pulses in connection with laser fusion experiments / 30 / or for the generation of X-band microwave pulses /31,32/. This article does not deal with high voltage applications.

In the next chapter technology and operation of picosecond optoelectronic switches will be described.

The following chapters give an overview on some interesting low voltage level applications with special regard to experiments performed by the author and his coworkers concerning the direct modulation of laser diodes and the generation of high-bit-rate electrical codes by picosecond optoelectronic switches.

---

* Experimental work presented here in part was performed at the Max-Planck-Institut für Festkörperforschung, Stuttgart, FRG, under financial support of the Max-Planck-Gesellschaft (MPG) and the Deutsche Forschungsgemeinschaft (DFG) and in part at the Technische Universität Wien, IAEE, Vienna, Austria, under financial support of the Austrian "Fonds zur Förderung der wissenschaftlichen Forschung (FWF)" under project no. S22/10.

* Paper presented at the AGARD-AVP specialists' meeting on "DIGITAL OPTICAL CIRCUIT TECHNOLOGY", Schliersee, Fed. Rep. Germany, September 1984

DISCUSSION

R.MacPherson, Ca

On your measurement diagram of the 100 photons per point and the several thousand points there seems to be some sort of three-fold symmetry in that pattern. Is there any significance in that?

Author's Reply

I have noticed the three-fold symmetry and haven't had time to work out why that exists yet — the results are very recent. The honest answer is, I don't know. It could at the moment be system imperfections because we are by no means anywhere near the quantum limit there. I suspect that's the answer, but I don't know. I'm not ignoring it — I just don't know the answer.

J.W.Haus, US

Do you expect to be able to measure squeezed states with this apparatus, and how much is the squeezed signal then degraded by the apparatus itself? Can you measure minimum uncertainty states?

Author's Reply

Certainly this sort of measurement scheme would be very sensitive towards squeezed states. We can imagine the noise distribution in an x-y measurement would become asymmetrical and that would come up very clearly as a sort of non-zero correlation in x and y measurements. So the answer is yes, it would be sensitive to that sort of measurement. As far as degraded is concerned, theoretically it's only degraded according to the quantum limit on determining what measurement you want to make on it; so in some sense no, but then practically it can be more difficult.

J.W.Haus, US

Well suppose you have a squeezing factor of two in one direction as compared to the other. Would your apparatus be able to pick it up?

Author's Reply

Theoretically it could pick it up — because if you remember those diagrams where I just had a fixed phase and had a spread of points on the phase and amplitude diagram. That would come up as an ellipse.

J.W.Haus, US

Was that spread, though, consistent with minimum uncertainty of the state?

Author's Reply

The spread you saw there was a long way from the quantum limit. The noise performance of the experiment still needs to go a long way yet before we measure quantum noise itself. That's just to show our interest.

[17] Kimble, H.J., Dagenais, M., Mandel, L.: 'Photon Antibunching in Resonance Fluorescence', Phys. Rev. Lett., v 39, n 11, Sept. 1977, pp. 691-694

[18] Dagenais, M., Mandel, L.: 'Investigation of Two-Time Emissions from a Single Atom', Phys. Rev. A, v 18, n 5, Nov. 1978, pp. 2217-2228

[19] Short, R., Mandel, L.: 'Observation of Sub-Poissonian Photon Statistics', Phys. Rev. Lett., v 51, n 5, Aug. 1983, pp. 384-387

[20] Yuen, H.P.: 'Two-Photon Coherent States of the Radiation Field', Phys. Rev. A, v 13, n 6, June 1976, pp. 2226-2242

[21] Yuen, H.P., Shapiro, J.H.: 'Optical Communication with Two-Photon Coherent States - Part I: Quantum State Propagation and Quantum-Noise Reduction', IEEE Trans. Inform. Thry., v IT-24, n 6, Nov. 1978, pp. 657-668

[22] Helstrom, C.W.: 'Estimation of a Displacement Parameter of a Quantum System', Int. J. Theor. Phys., v 11, n 6, 1974, pp. 357-378

[23] Fedorov, S.YE., Mart'Yanov, A.N.: 'An Investigation of the Noise Immunity of Detection of Optical Coherent Signals with a Multilevel Phase Manipulation', Radio Eng. and Electron. Phys., v 28, n 2, 1983, pp. 59-64, (Translation from Russian)

[24] Carruthers, P., Nieto, M.: 'Phase and Angle Variables in Quantum Mechanics', Rev. Mod. Phys., v 40, n 2, Apr. 1968, pp. 411-440

Acknowledgements

This project is now funded by the Joint Optical and Electronic Research Scheme. We are indebt to Standard Telecommunications Laboratory Ltd. for some early funding. One of the authors (Mr. N. G. Walker) would like to thank the Science and Engineering Research Council (U.K.) for a personal grant.

Conclusions

We have shown how microwave multiport techniques may be extended to optical frequencies to realise potentially useful measurements, and are currently building a homodyne experiment which demonstrates these principles. The system has not yet been calibrated so figures on its noise performance are not available, however, we would like to obtain good enough performance to allow the measurement of quantum noise.

A quantum mechanical approach to optical multiports has been outlined, and some of the implications of this touched upon. In particular, we presented the quantum operator corresponding to the measurement made by a symmetric four power detection scheme aimed at determining the complex amplitude of a signal. This multiport measures the 'probability operator measure' characterised by a resolution of the identity into coherent states (equation 18). The importance of the quantum theory is that the noise statistics are determined by the quantum state of the weak signal field rather than the local oscillator shot noise and these could, therefore, be far from Gaussian or Poisson in the most general case. This could have important consequences on system performance if Shapiro or states with sub-Poissonian photon statistics, became available for use in real systems.

The performance of our multiport equivalent quantum operator in measuring the phase of a coherent state signal falls short of the quantum optimum since the multiport inherently measures both amplitude and phase whose quantum operators do not commute. This raises the question as to how the optimum phase measurement may be made, and suggests that a general analysis of optical multiports with one photon, two photon, and higher order detectors would be worthwhile.

References

[1] Engen, G.F., Hoer, C.A.: 'Application of an arbitrary 6-port junction to power measurement problems', IEEE Trans Instrum. Meas., v IM-21, 1972, pp. 470-474

[2] Engen, G.F.: 'The Six-Port Reflectometer: An Alternative Network Analyser', IEEE Trans Microwave. Thry. Techn., v MTT-21, n 12, 1977, pp. 1075-1079

[3] Luff, G.F., Probert, P.J., Carroll, J.E.: 'Real-Time Six-Port Reflectometer', IEE Proc. H, v 131, 1984, pp. 186-190

[4] See for example, Ramo, S., Whinnery, J., Van Duzer, T.: 'Fields and Waves in Communication Electronics', John Wiley & Sons., Inc., 1965, pp. 627-631

[5] Burns, W.K., Moeller, R.P., Villarruel, C.A.: 'Observation of Low Noise in a Passive Fibre Gyroscope', Electronics Letters, v 18, n 15, July 1982, pp. 648-650

[6] Helstrom, C.W.: 'Quantum Detection and Estimation Theory', Academic Press, Inc., 1976

[7] Helstrom, C.W., Liu, J.W.S., Gordon, J.P.: 'Quantum-Mechanical Communication Theory', Proc. IEEE, v 58, n 10, Oct. 1970, pp. 1578-1598

[8] Dolinar, S.J.: 'An Optimum Receiver for the Binary Coherent State Quantum Channel', Res. Lab. Electron. M.I.T. Camb. Quart. Prog. Rep., n 111, Oct. 1973, pp. 115-120

[9] Kennedy, R.S.: 'A Near-Optimum Receiver for the Binary Coherent State Quantum Channel', Res. Lab. Electron. M.I.T. Camb. Quart. Prog. Rep., n 108, Jan 1973, pp. 219-225

[10] Shapiro, J.H.: 'On the Near-Optimum Binary Coherent State Receiver', IEEE Trans. Inform. Thry., v IT-26, n 4, July 1980, pp. 490-491

[11] Yuen, H.P., Shapiro, J.H.: 'Optical Communication with Two-Photon Coherent States - Part III: Quantum Measurements Realizable with Photoemissive Detectors', IEEE Trans. Inform. Thry., v IT-26, n 1, Jan. 1980, pp. 78-92

[12] Kelly, P.L., Kleiner, W.H.: 'Theory of Electromagnetic Field Measurement and Photoelectron Counting', Phys. Rev. A, v 136, n 2A, Oct. 1964, pp. A316-A334

[13] Yuen, H.P., Shapiro, J.H.: 'Optical Communication with Two-Photon Coherent States - Part II: Photoemissive Detection and Structured Receiver Performance', IEEE Trans. Inform. Thry., v IT-25, n 2, March 1979, pp. 179-192

[14] Louden, R.: 'Non-Classical effects in the Statistical Properties of Light', Rep. Prog. Phys., v 43, 1980, pp. 913-949, (Review)

[15] Paul, H.: 'Photon Antibunching', Rev. Mod. Phys., v 54, n 4, Oct 1982, pp. 1061-1102, (Review)

[16] Walls, D.F.: 'Squeezed States of Light', Nature, v 306, 10th Nov. 1983, pp. 141-146, (Review)

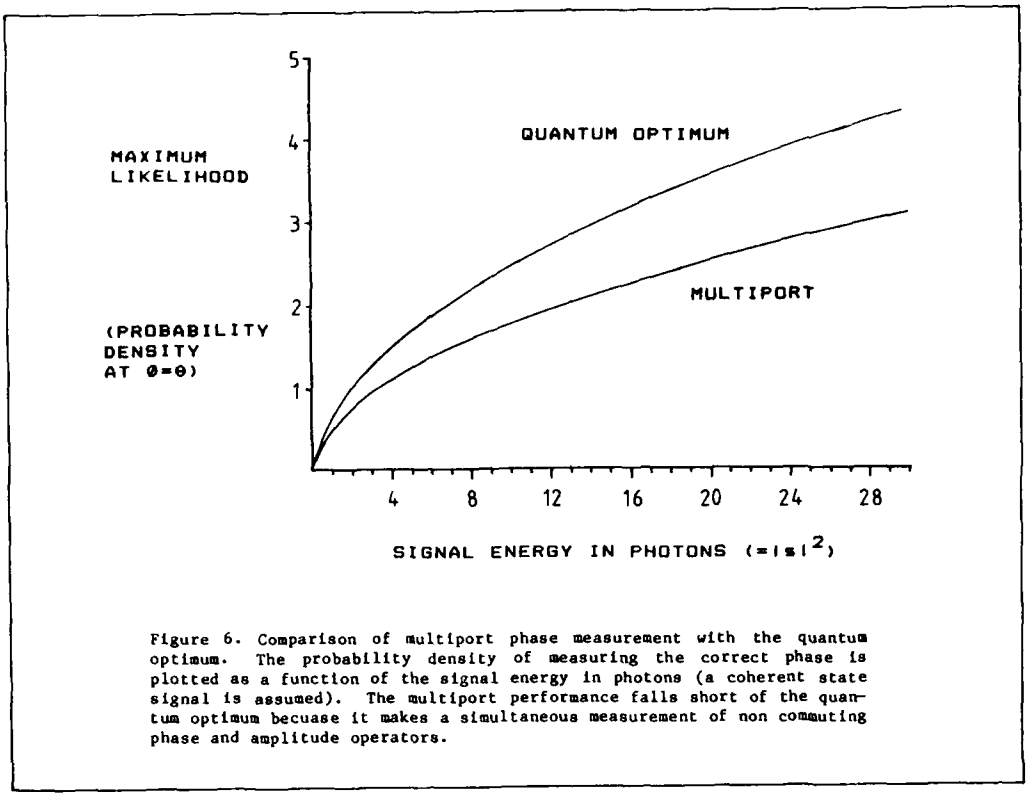


Figure 6. Comparison of multiport phase measurement with the quantum optimum. The probability density of measuring the correct phase is plotted as a function of the signal energy in photons (a coherent state signal is assumed). The multiport performance falls short of the quantum optimum becuase it makes a simultaneous measurement of non commuting phase and amplitude operators.

The performance of our multiport in measuring the phase of a coherent signal can be compared to that of the quantum optimum estimator. The probability operator measure for optimum phase estimation is characterised by the resolution of the identity [6;pp. 243-248,22,23]

$$\int \frac{|\phi\rangle\langle\phi|}{2\pi} d\phi = 1 \qquad (28)$$

where

$$|\phi\rangle = \sum_{n=0}^{\infty} \exp(jn\phi)\ |n\rangle \qquad (29)$$

and $|n$ is a number state with n photons. We call $|\phi\rangle$ the phase states. The probability density for phase measurement is then

$$\text{Prob}(\phi) = \frac{|\langle\phi|s\rangle|^2}{2\pi} \qquad (30)$$

Putting s real, writing the coherent state $|s\rangle$ as a superposition of number states and using the orthogonality property of these gives

$$\text{Prob}(\phi) = \frac{1}{2\pi} \left| \sum_{1}^{\infty} \frac{r^n}{\sqrt{n!}} \right|^2 \exp(-|r|^2) \qquad (31)$$

This is also plotted in figure 6.

The multiport phase measurement performance falls short of the quantum optimum, however, no scheme has yet been suggested for realising quantum optimum phase measurements. The basic reason for the sub-optimum performance is that a multiport inherently gives both amplitude and phase information about the signal, but the quantum phase and amplitude operators do not commute [24]. By measuring both variables simultaneously one introduces excess noise in each measurement. In order to realise the optimum phase measurement it is necessary to use the full quantum theory of the field, rather than the semiclassical shot noise detection model, in order to exploit the properties of the extra dimensionality thereby introduced. There is more interesting work to be done on this problem.

$$\frac{1}{2}\left[\frac{1}{k} - 1\right] \qquad (21)$$

As the quantum efficiency of the detectors reduces so the quantum statistics of the signal state become masked in Gaussian noise, and the signal state can not be distinguished from a coherent state. This turns out to be true for all multiport measurements. We continue to consider the ideal case with k = 1.

It can be shown that if the signal is in a coherent state or a random superposition of coherent states, then the noise behaviour is exactly the same as that predicted by the semiclassical Poisson process model. However there exist states of the quantised field which have no classical analog [14,15,16] and in this case the semiclassical detection model no longer gives the correct noise statistics. There has been a lot of literature published recently on possible schemes for generating non-classical states, although only two experimental sightings have been made so far [17,18,19]. The attraction of non-classical states is the possibility of noise reduction. For example the two photon coherent states discussed by Yuen [20] and Shapiro [21,13,11] have reduced uncertainty in an 'in phase' measurement at the expense of increased uncertainty in the 'quadrature' component. There is also interest in states which have sub-Poisson photon counting statistics (with variance less than the mean) and hence reduced noise in a direct detection application.

The fully quantised model of multiports is not only essential for correctly describing the detection statistics of non-classical quantum states, but also provides a more attractive representation of the homodyne measurements than the semiclassical model.

Phase Measurements

We are particularly interested in phase measurements using optical multiports and their possible application to optical communication systems. In a two level phase shift keying system a decision has to be made only between two alternatives; complete knowledge of the phase of the signal is not necessary provided the local oscillator phase is locked to the signal zero phase. In this case a single beamsplitter will form a suitable mixer for the signal and local oscillator. However, if the actual value of phase has to be estimated (in an analogue sense) then at least three power measurements are required. The detection scheme we have outlined here uses four measurements. To evaluate the performance of the multiport in measuring phase we can choose any of the criterion of maximum likelihood, minimum variance, or any other Baye's cost function. We consider the likelihood of measuring the correct phase value, or in other words, the value of the signal phase measurement probability density function at the value of phase originally sent by the transmitter.

Our multiport measures the probability operator measure characterised by the resolution of the identity

$$\int \frac{|r\rangle\langle r|}{\pi} d^2r = 1 \qquad (22)$$

So that if the signal field is in state $|s\rangle$ then the probability density function of measuring the in phase and quadrature components to be r_1 and r_2 respectively is

$$\text{Prob}(r_1,r_2) = \frac{|\langle r|s\rangle|^2}{\pi} \qquad (23)$$

where $r = r_1 + jr_2$. If $|s\rangle$ is a coherent state then this becomes

$$\text{Prob}(r_1,r_2) = \frac{1}{\pi}\exp(-|r-s|^2) \qquad (24)$$

Since we are interested in measuring the phase of the signal, but not its amplitude we transform to polar coordinates.

$$r = |r|\exp(j\phi) \quad \text{and} \quad s = |s|\exp(j\theta) \qquad (25)$$

θ is the phase of the signal and ϕ is our multiport estimate of θ. The probability distribution for ϕ and $|r|$ is

$$\text{Prob}(|r|,\phi) = \frac{1}{\pi}|r|\exp\left[-(|r|^2 + |s|^2 - 2|r||s|\cos(\phi-\theta))\right] \qquad (26)$$

To obtain the probability distribution for a measurement of ϕ alone this is integrated over $|r|$. The probability density is greatest when $\phi = \theta$ and the estimate is unbiased. Putting $\phi = \theta$ and doing the integration gives

$$\text{Prob}(\phi = \theta) = \left[\frac{|s|^2}{\pi}\right]^{1/2}\left[\frac{1}{2}\left[1 + \text{erf}(|s|)\right] + \frac{1}{\sqrt{4\pi|r|^2}}\exp(-|s|^2)\right] \qquad (27)$$

The value of this maximum likelihood estimate is plotted as a function of $|s|^2$ in figure 6. $|s|^2$ is the average number of photons in a coherent state $|s\rangle$.

The statistics of the photoemissions from all the detectors are then identical to those of a simultaneous measurement of all the $\hat{N}_m$. A complete analysis must include the simultaneous measurements to allow for the possibility of correlation between different ports. The state of the input fields is described by their density operator, which is assumed to be the tensor product of all the separate density operators for each input port field.

We will not give a detailed derivation of the multiport photoelectron counting statistics, but will present the results of one example using the four power measurement multiport. We investigate homodyne detection with a strong local oscillator, entering port a, so that this dominates over any of the other signals into the multiport. We have one input into port b which is our unknown signal. The signal field may be in any quantum state and we wish to estimate its complex amplitude. All other inputs to the multiport are assumed to be in the coherent ground state or, equivalently, the zero number state.

As before the scattering matrix can be split into a forward half and reverse half. Furthermore we adopt similar processing to equations 10 and 11 and take the differences in the outputs

$$r_1 = \frac{1}{|a_a|}(P_e - P_f) \quad \text{and} \quad r_2 = \frac{1}{|a_a|}(P_g - P_h) \tag{15}$$

The algebra in deriving the measurement statistics for r_1 and r_2 is rather long, but follows a similar line to the analysis of homodyning using a single beamsplitter but with multimode fields given by Yuen and Shapiro [11]. It will not be repeated here.

The result is that the multiport measures the probability operator measure characterised by the resolution of the identity into coherent states (for further explanation of this see for example [6] pp. 53-55 and pp. 128-133) In other words if the signal field is in state $|s\rangle$ then the probability of measuring the complex amplitude to be r_1+jr_2 is

$$\begin{aligned} \text{Prob}(r_1,r_2) &= \frac{1}{\pi}\langle s|r_1+jr_2\rangle\langle r_1+jr_2|s\rangle \\ &= \frac{1}{\pi}|\langle r_1+jr_2|s\rangle|^2 \end{aligned} \tag{16}$$

where $|r_1+jr_2\rangle$ is a coherent state with

$$\hat{a}\,|r_1+jr_2\rangle = (r_1+jr_2)\,|r_1+jr_2\rangle \tag{17}$$

The coherent state resolution of the identity is

$$\int \frac{|r\rangle\langle r|}{\pi}\, d^2r = 1 \tag{18}$$

where the r integration is over the entire plane (r_1,r_2)

Using this we have immediately

$$\int \text{Prob}(r_1,r_2)\, dr_1\, dr_2 = 1 \tag{19}$$

as must be the case for a probability distribution. We see that, for our multiport, the projection of the system state onto the coherent state of given complex amplitude gives the probability of measuring the signal to have that complex amplitude.

Thus the statistical nature of measurements made by our symmetric four power measurement multiport have been described in purely abstract quantum terms, and this characterisation holds regardless of the state of the quantum signal entering the measurement port b. In fact our multiport homodyning measures the same operator as a conventional heterodyne receiver [11]. It is interesting to note that in this model the noise is due to a quantum measurement of the signal field rather than the local oscillator shot noise. We thus have to consider rather carefully the signal state when evaluating the performance of a detection system as the noise in the detectors is not in general uncorrelated or Gaussian.

If the detectors have sub-unity quantum efficiency then we have to take this into account in the estimate of the incident signal amplitude. If the estimate of the complex amplitude is r_1+jr_2 and the quantum efficiency is k then we must adopt the processing

$$r_1 = \frac{1}{k|a_a|}(P_e - P_f) \quad \text{and} \quad r_2 = \frac{1}{k|a_a|}(P_g - P_h) \tag{20}$$

The noise statistics of this multiport measurement are exactly the same as with the unit quantum efficiency case except with superposed classical Gaussian additive noise of variance

An integration of both devices - the GaAs:Cr optoelectronic switch and GaAs/GaAlAs heterostructure laser - should be possible on a common (GaAs-) substrate. A fast electrical modulator wich can be driven externally without requiring high-speed cable connections and a fast light emitting device thus could be integrated monolithically. This combination could be realized also using other semiconductor materials (e.g. InP:Fe switch and InGaAsP/InP laser diodes) and may be of interest for high speed integrated optical circuit technology.
If the ultrafast driven laser diode of figure 7 is located within an external cavity, matched in length to the pump pulse sequence, a synchronous mode-locked operation of the semiconductor laser can be achieved (figure 9) /45/.

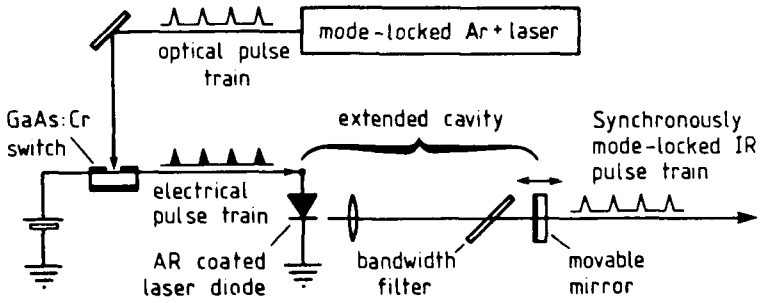


Figure 9
Synchronous mode-locking of a semiconductor laser by a picosecond optoelectronic switch

As pump laser for the optoelectronic switch can be used an actively mode-locked Ar^+ ion laser with a typical repetition rate of 80 MHz and single pulse widths of ≈ 100 ps. With a commercially available GaAs/GaAlAs buried heterostructurelaser diode which was anti-reflection coated on one laser facet output pulses of 30 ps duration (FWHM) have been obtained strongly synchronized to the pulse train of the Ar^+ ion laser pump. The laser emission spectrum tained during this short pulse regime, however, has been multimode and only by inserting a 80 µm thick Fabry-Perot etalon (R=30 %) into the laser cavity single-mode operation can be achieved. Under these conditions the pulse width of the mode-locked pulse has been 45 ps; the spectral width of one longitudinal mode has been measured to be 25 GHz which results in a time bandwidth product of 1.1 (figure 10).

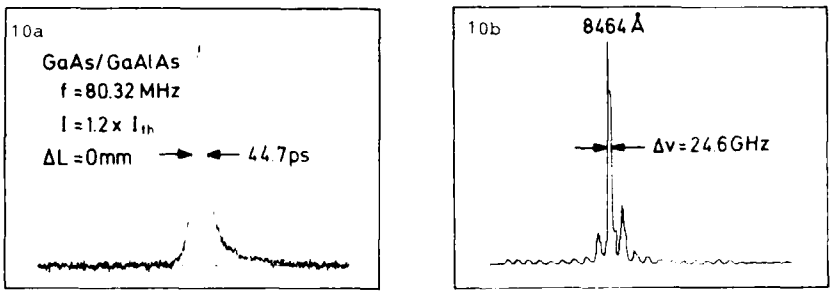


Figure 10 : Streak camera trace (10a) and spectrum (10b) of a GaAs/GaAlAs BH-laser synchronously pumped by a GaAs:Cr optoelectronic switch /45/.

Since the mode-locked pump laser could simultanuously also be used to synchronously mode-locking of dye-lasers the spectrals range accessible to "probe and exite" picosecond experiments could be extended to the infrared region wherever semiconductor lasers are available.
By the use of optical delay lines the pump pulse sequences can be increased to the GHz range, whereby the external cavity length of the laser diode has to be in the order of several centimeters, which again could be realized on integrated optical chips. The output pulse train of the mode-locked external cavity laser has an extremely good temporal stability and can e.g. be used as internal clock sequence for Gbit/s optical computation systems.

7. HIGH-BIT-RATE ELECTRICAL CODE GENERATION

Gbit/s code generators are of great interest for high-bit-rate communication and sensing systems and for testing of GHz electronic instrumentation /51/. Different approaches for the electrical bit-pattern generation in the Gbit/s range have been reported: digital electronic techniques /52/, optical fiber tapped delay line techniques /53/, and surface acoustic wave tapped delay line techniques /54/. We have used an optoelectronic technique for electrical pulse code generation at a bandwidth ≥ 10 GHz. Our technique is based on the optoelectronic transformation of a high-bit-rate optical to a corresponding electrical pulse code by means of a picosecond photoconductive detector. The high-bit-rate optical code is obtained by sending a picosecond pulse through a multiple beam path optical delay line (MOD) (see Fig. 11). The different beams, after having suffered different delays, are recombined by a single focusing lens (f=30 mm) onto the 30 µm microstrip-gap of a picosecond optoelectronic switch. The photoactive material of the detector is a 1 µm thick silicon-on-sapphire layer which has been irradiated by an appropriate dose of 50-200 keV O^+ ions to obtain a high density of trapping defects for the optically generated carriers /17/.

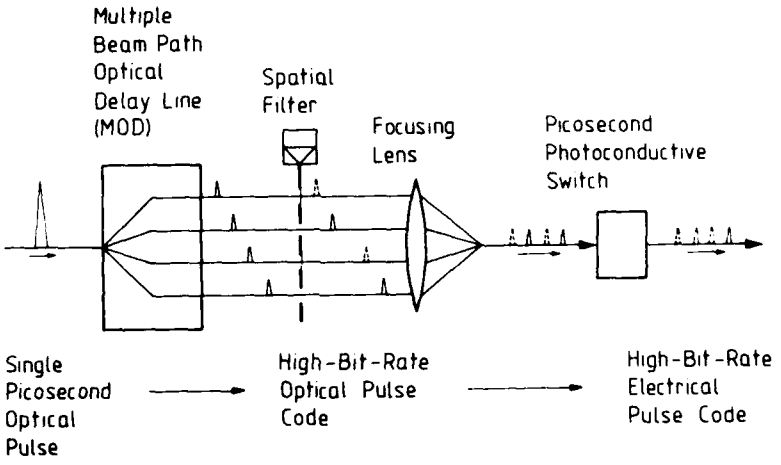


Figure 11 : Principle diagram for the optoelectronic high-bit-rate code generation using a picosecond photoconductive switch.

The response time of the SOS photoconductive device has been determined by an optoelectronic autocorrelation measurement(see chapter 4) to be 18 ps (FWHM).

Figure 12 a shows the response of the SOS-photoconducting switch if activated by a pair of optical picosecond pulses (pulse with 5 ps FWHM, wavelength 720 mm) which are separated by a time interval of 70 ps. The two pulses are clearly resolved by the sampling head unit. The pulse sequence corresponds to a bandwidth of 14.3 GHz. The delay between the two pulses has been obtained after the passing of the picosecond laser beam through a delay line scheme standardly used for nonlinear optical autocorrelation measurements. If each of the two parallel beams is split again in two parts, four identical parallel beams can be obtained which can be delayed deliberately with repect to each other. If more (n) beams are needed (for the formation of a n-bit word) the arrangement can be extended to a multiple (n) beam path optical delay line (MOD). The MOD can deliver any n-bit optical code at bit rates up to 100 GHz if each single beam path can be blocked separately by a spatial filter. The MOD which in our case consisted of conventional macroscopic optical components can be replaced by an array of fiber optic delay lines.

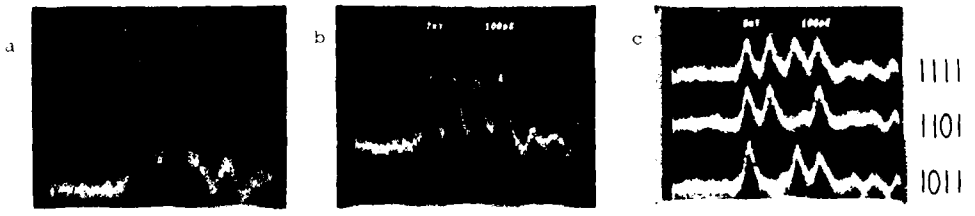
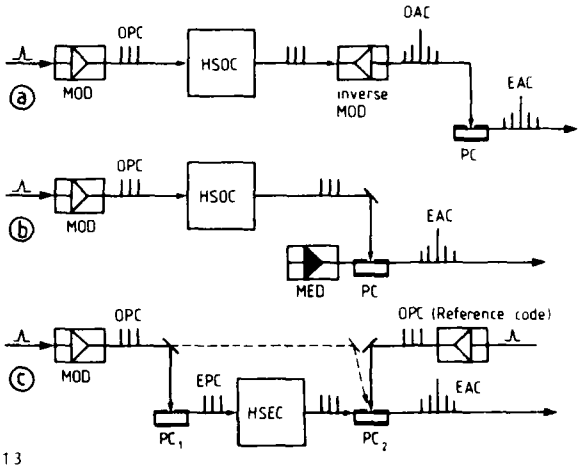


Figure 12
Sampling oscilloscope traces of high-bit-rate electrical pulses generated by a SOS picosecond photoconductive switch:
a) Response of SOS device, activated by a pair of picosecond optical pulses with a time delay of 70 ps corresponding to a bandwidth of 14 GHz.
b) Response of SOS device activated by four optical picosecond pulses with a time delay of 100 ps (bandwidth 10 GHz)
c) Electrical 4-bit code generation at 10 Gbit/s (code words: 1111, 1101, 1011)

In Figure 12 b,c different 4-bit electrical code words (Fig.12b: 1111; Fig.12c: 1111,1010, 0101)corresponding to transmission rates of10 Gbit/s are shown. The optical code pulse trains have been generated by introducing a spatial filter into the 4-beam-MOD.
The ultimate transmission rate of the system is limited by the speed by which the spatial filtering of the multiple beam array can be actively manipulated. An integrated optical system containing planar electrooptic or acoustooptic Bragg deflectors could be able to control the spatial filtering at rates up to ≈ 1 GHz /55/. This should lead to an optical communication system which transmits n-bit code words (bandwidth 10 GHz) at a duty cycle of ≈ 1 GHz. The high-bit-rate electrical pulse train could be used to directly modulate a semiconductor laser /12,44/ as a source of a Gbit/s optical communication system.
Different schemes could be employed for matched filtering and decoding of the high-bit-rate optical or electrical pulse sequence (see figure 13). Optical matched filtering can be obtained when the high-bit-rate optical pulse train is inserted inversely into the MOD /53/ (figure 13a). Electrical matched filtering of the Gbit/s electrical pulse code could be done by feeding it into a n-path electrical microstrip delay line circuit (figure 13b). Decoding of the high-bit-rate electrical code could be performed optoelectronically if a second ultrafast photoconductive detector activated by an appropriate optical pulse sequence is used as a sampling gate /56/ (figure 13c)



a) Optical matched filtering by inverse MOD.

b) Decoding of high-bit-rate optical codes OPC by electrical mached filtering in a multiple path electrical microstrip delay line (MED)

c) Decoding of high-bit-rate electrical codes (EPC) by high-speed optoelectronic sampling with an optical reference code source.

Figure 13
Possible schemes for matched filtering and decoding of single high-bit-rate optical (OPC) or electrical (EPC) pulse codes.
(Abbreviations: MOD...multiple-beam-path optical delay line; MED...multiple-path electrical microstrip delay line; HSOC...high-speed optical circuit; HSEC...high-speed electrical circuit ; OAC...optical autocorrelation; EAC...electrical autocorrelation; PC... picosecond photoconductive device).

8. CONCLUSION

Picosecond optical electronics which introduces the high-speed capability of picosecond laser technology to microelectronic instrumentation has become during recent years one of the most important areas of application for ultrashort laser pulses /57/. The author of this contribution by far could not mention all publications and recent developments on this rapidly expanding field. For example, of considerable technical importance, which could not be treated in detail, is the high voltage switching capability of picosecond photoconductors which has been used, e.g. to drive Pockels cells for prepulse suppression in laser fusion experiments /30/ or to trigger jitter-free a streak camera for time resolved picosecond spectroscopy /29/ or for the generation of high power microwave pulses /31,32/. For a review on this topic the reader is referred to /26/.
The particular examples discussed in this article are chosen to give a basic review on high speed photoconductive switching by picosecond laser pulses to illustrate some of the low voltage level applications of picosecond photoconductors for Gbit/s optical circuit technology. The new type photoconductive switches can be used as high speed photodetectors as well as ultrafast switches and sampling gates for the characterisation of high speed optoelectronic materials and devices. Furthermore they can be considered as sources for ultrashort electrical pulses and high-bit-rate electrical codes ,which is of special importance for the development of high-speed optical communication and optical computing systems. Some fundamental experiments concerning this topic have been performed by the author and his coworkers and are discussed more in detail in this article. The possibility of a monolithical integration of picosecond photoconductors and laser diodes on high speed semiconductor microcircuits should be emphasized. The ultimate speed limits of picosecond optical electronics have not been reached up to now and with a further optimization of materials and high frequency circuit design the achievement of the subpicosecond range seems to be realistic. The direct transformation of extremely high-bit-rate optical pulse trains to corresponding electrical pulse sequences by means of picosecond optoelectronic switching opens the way to multi-gigabit-per-second optical communication and optical computing systems which presents a challenge for the optical circuit technology of the next future.

REFERENCES

1 D.H. Auston, "Picosecond optoelectronic switching and gating in silicon", Appl. Phys. Lett., vol. 26, 1975, pp. 101 - 103
2 A.M. Johnson, D.H. Auston, "Microwave switching by picosecond photoconductivity", IEEE J. Quantum Electron., vol. QE-11, 1975, pp. 283 - 287
3 D.H. Auston, P. Lavallard, N. Sol, D. Kaplan, "An amorphous silicon photodetector for picosecond pulses", Appl. Phys. Lett., vol. 36, 1980, pp. 66 - 68
4 D.H. Auston, A.M. Johnson, P.R. Smith, J.C. Bean, "Picosecond optoelectronic detection, sampling, and correlation measurements in amorphous semiconductors", Appl. Phys. Lett., vol. 37, 1980, pp. 371 - 373
5 A.M. Johnson, D.H. Auston, P.R. Smith, J.C. Bean, J.P. Harbison, D. Kaplan, "Picosecond photoconductivity in amorphous silicon", Proceedings of the Second Topical Meeting on Picosecond Phenomena, Cape Cod (USA), 1980, pp. 285 - 289
6 A.M. Johnson, D.H. Auston, P.R. Smith, J.C. Bean, J.P. Harbison, A.C. Adams, "Picosecond transient photocurrents in amorphous silicon", Phys. Rev. B, vol. 23, 1981, pp. 6816 - 6818
7 A.P. DeFonzo, "Picosecond photoconductivity in germanium films", Appl. Phys. Lett., vol. 39, 1981, pp. 480 - 482
8 R.A. Lawton, A. Scavannec, "Photoconductive detector of fast-transition optical waveforms", Electron. Lett., vol. 11, 1975, pp. 74 - 75
9 C.H. Lee, "Picosecond optoelectronic switching in GaAs", Appl. Phys. Lett., vol. 30, 1977, pp. 84 - 86
10 G. Mourou, W. Knox, "High-power switching with picosecond precision", Appl. Phys. Lett., vol. 35, 1979, pp. 492 - 495
11 G. Veith, A. Perger, K. Lübke, A.J. Schmidt, unpublished measurements, TU Wien, Vienna 1981
12 E.O. Göbel, G. Veith, J. Kuhl, H.-U. Habermeier, K. Lübke, A. Perger, "Direct gain modulation of a semiconductor laser by a GaAs picosecond optoelectronic switch", Appl. Phys. Lett., vol. 42, 1983, pp. 25 - 27
13 F.J. Leonberger, P.F. Moulton, "High speed InP optoelectronic switch", Appl. Phys. Lett., vol. 35, 1979, pp. 712 - 714
14 J. Degani, R.F. Leheny, R.E. Nahory, M.A. Pollak, J.P. Heritage, J.C. DeWinter, "Fast photoconductive detector using p-$In_{0.53}Ga_{0.47}As$ with response to 1.7 μm", Appl. Phys. Lett., vol. 38, 1981, pp. 27 - 29
15 A.G. Foyt, F.J. Leonberger, R.C. Williamson, "Picosecond InP optoelectronic switches", Appl. Phys. Lett., vol. 40, 1982, pp. 447
16 P.M. Downey, D.H. Auston, P.R. Smith, "Picosecond correlation measurements of indium phosphide photoconductors", Appl. Phys. Lett., vol. 42, 1983, pp. 215 - 217
17 P.R. Smith, D.H. Auston, A.M. Johnson, W.M. Augustyniak, "Picosecond photoconductivity in radiation-damaged silicon-on-sapphire films", Appl. Phys. Lett., vol. 38, 1981, pp. 47 - 50
18 T. Miyoshi, P.K. Tien, R.J. Martin, D.M. Tennant, A.M. Johnson, P.M. Downey, "Infrared photodetection in proton-bombarded InP", Appl. Phys. Lett., vol. 44, 1984, pp. 128 - 130
19 P.M. Downey, B. Schwartz, "Picosecond photoresponse in $^3He^+$ bombarded InP photoconductors", Appl. Phys. Lett., vol. 44, 1984, pp. 207 - 209
20 R.B. Hammond, N.G. Paultier, R.S. Wagner, T.E. Springer, "Excitation and Fe concentration dependences in the impulse photoconductance of InP:Fe", Appl. Phys. Lett., vol. 44, 1984, pp. 620 622
21 V. Brückner, F. Kerstan, "One and two-photon excited picosecond photoconductivity in CdS", Opt. Commun., vol. 45, 1983, pp. 187 - 192
22 W. Margulis, W. Sibbett, "Picosecond CdSe photodector", Appl. Phys. Lett., vol. 42, 1983, pp. 975 - 977
23 See e.g.: S.L. Shapiro (ed.), "Ultrashort light pulses", Topics in Appl. Phys., vol. 18, Springer Verlag 1977
24 R.L. Fork, C.V. Shank, R. Yen, C.A. Hirliman, "Femtosecond optical pulses", IEEE J. Quantum Electron., vol. QE-19, 1983, pp. 500 - 505
25 D.H. Auston, P.R. Smith, "Picosecond optical electronics for high-speed instrumentation", Laser Focus, vol. 18, April 1982, pp. 89 - 93
26 G. Mourou, W. Knox, S. Williamson, "Picosecond high-power switching and applications", Laser Focus, vol. 18, April 1982, pp. 97 - 106
27 P.R. Smith, D.H. Auston, W.M. Augustyniak, "Measurement of GaAs field-effect transistor electronic impulse response by picosecond optical electronics", Appl. Phys. Lett., vol. 39, 1981, pp. 739 - 741
28 D.H. Auston, P.R. Smith, "Picosecond optical electronic sampling: Characterization of high-speed photodetectors", Appl. Phys. Lett., vol. 41, 1982, pp. 599 - 601
29 M. Stavola, G. Mourou, W. Knox, "Picosecond time delay fluorimetry using jitter-free streak camera", Opt. Comm., vol. 34, 1980, pp. 404 - 408
30 G. Mourou, J. Bunkenburg, W. Seka, "Electrooptic prepulse suppression for fusion laser systems", Opt. Commun., vol. 34, 1980, pp. 252 - 254
31 G. Mourou, C.V. Stancampiano, D. Blumenthal, "Picosecond microwave pulse generation", Appl. Phys. Lett., vol. 38, 1981, pp. 470 - 472
32 C.H. Lee, S. Mak, A.P. DeFonzo, "Millimetre-wave switching by optically generated plasma in silicon", Electron. Lett., vol. 14, 1978, pp. 733 - 734
33 C.H. Lee, V.K. Mathur, "Picosecond photoconductivity and its applications", IEEE J. Quantum Electron., QE-17, pp. 2098 - 2112
34 D.H. Auston, "Impulse response of photoconductors in transmission lines", IEEE J. Quantum Electron., vol. QE-19, 1983, pp. 639 - 648
35 B.-Y. Hwang, F.A. Lindholm, R.B. Hammond, "Methodology for calculating turn-off transient of photoconductivity circuit elements in picosecond optoelectronics", IEEE J. Quantum Electron., vol. QE-19, 1983, pp. 648 - 657

36 Ch. Kittel, "Introduction to Solid State Physics", 5th edition, John Wiley & Sons, Inc., New York, 1976, pp. 230 - 231

37 M. Stavola, M.G. Sceats, G. Mourou, "Picosecond switching of a multi-kilovolt dc bias with laser activated silicon at low temperature", Opt. Commun., vol. 34, 1980, pp. 409 - 412

38 G. Mourou, W. Knox, "A picosecond jitter streak camera", Appl. Phys. Lett., vol. 36, 1980, pp. 623 - 626

39 J. Bokor, P.H. Bucksbaum, D.H. Auston, "Measurement of picosecond ultraviolet pulse-widths using an electrical autocorrelator", Appl. Phys. Lett., vol. 42, 1983, pp. 342 - 344

40 D.H. Auston, R.R. Freeman, P.R. Smith, D.M. Mills, R.H. Siegmann, "High speed x-ray sensitive photoconducting detector", Appl. Phys. Lett., vol 42, 1983, pp. 1050 - 1052

41 G. Veith, E.O. Göbel, J. Kuhl, "Optoelectronic 10 Gbit/s code generation by a picosecond photoconductive device", Electron. Lett., vol. 19, 1983, pp. 1072 - 1073

42 W. Margulis, W. Sibbett, J.R. Taylor, "Active mode- locking of lasers using GaAs and GaP picosecond switches", Opt. Commun., vol. 35, 1980, pp. 153 - 156

43 J.A. Valdmanis, G. Mourou, C.W. Gabel, "Subpicosecond electrical sampling", IEEE J. Quantum Electron., QE-19, 1983, pp. 664 - 667

44 G. Veith, J. Kuhl, E.O. Göbel, "Temporal and spectral characteristics of rapidly gain-switched GaAs/GaAlAs burried-heterostructure lasers", Electron. Lett., vol. 19, 1983, pp. 385 - 386

45 E.O. Göbel, J. Kuhl, G. Veith, "Synchronous mode-locking of semiconductor laser diodes by a picosecond optoelectronic switch", J. Appl. Phys., vol. 56, 1984, pp. 862-864

46 G. Veith, "Picosecond optoelectronics - High speed optoelectronic switching by picosecond laser pulses", Laser & Optoelektronik, vol. 15, no. 2, 1983, pp. 117 - 125

47 K.K. Li, G. Arjavalingam, A. Dienes, J.R. Whinnery,"Propagation of picosecond pulses on microwave striplines", IEEE MTT, vol. 30, 1982, pp. 1270 - 1273

48 E.O. Göbel, J. Kuhl, G. Veith, "Pulse modulation of double heterostructure diode lasers by picosecond optoelectronic switches", Proc. SPIE Picosecond Optoelectronics, San Diego, USA, Aug. 1983, pp. 79 - 86

49 T. Andersson, S. Lundquist, S.T. Eng, "Generation of single mode picosecond pulses by injection locking of an AlGaAs semiconductor laser", Appl. Phys. Lett., vol. 41, 1982, pp. 14 - 16

50 G. Arnold, P. Russer, K. Petermann, "Modulation of laser diodes", in "Semiconductor Devices for Optical Communications", ed. H. Kressel, Topics in Appl. Phys., vol. 39, Springer Verlag 1980, pp. 213 - 242

51 B.G. Bosch "Gigabit Electronics - A Review", Proc. IEEE, vol. 67, 1979, pp. 340-379

52 C.A. Liechti, G.L. Baldwin, E. Cowen, M. Namjoo,A. Podell "A GaAs word generator operating at 5 Gbit/s data rate".
IEEE Trans. MTT, vol. MTT-30, 1981, pp. 998-1006 (Conference paper given at the 1981 IEEE GaAs Integrated Circuits Symposium, San Diego 1981)

53 K.P. Jackson, S.A. Newton, H.J. Shaw, "1 Gbit/s code generator and matched filter using an optical fiber tapped delay line", Appl. Phys. Lett., vol 42, 1983, pp. 556-558.

54 P.H. Carr, R.D. Colvin, J.H. Silva, J. Deklerk, B.R. McAvoy,"Encoding and Decoding at 1 GHz with SAW tapped delay lines".
IEEE Ultrasonics Symposium Proceedings, New Orleans 1979, pp. 757-760

55 F. Auracher " Planar electrooptic ans acoustooptic Bragg- deflectors" in "Integrated Optics" ed. by S. Martellucci and A.N. Chester; Plenum Publishing Corporation 1983, pp. 211-223.

56 D.H. Auston, P.R. Smith "Generation and Detection of Millimeter Waves by Picosecond Photoconductivity" Appl. Phys. Lett. vol 43, 1983, pp. 631-633.

57 see for example: "Picosecond Phenomena III" ed. by K.B. Eisenthal, R.M. Hochstrasser, W. Kaiser, A. Laubereau, Proc. of the 3th Internat. Conf. on Picosecond Phenomena, Garmisch Partenkirchen (FRG), 1982 (Springer Series in Chemical Physics, vol. 23, Springer Verlag Berlin Heidelberg, New York, 1982)

ACKNOWLEDGEMENT

The author would like to thank E.O. Göbel and J. Kuhl for the fruitful cooperation during his stay at the Max-Planck- Institut für Festkörperforschung, Stuttgart (FRG) which was financially supported by the Max-Planck-Gesellschaft (MPG).
He gratefully acknowledges also cooperation with A.J. Schmidt, A. Perger and K. Lübke during an earlier stage of his work at the Technische Universität Wien, IAEE, Vienna (Austria) and financial support by the Austrian science research fund (FWF).

DISCUSSION

S.Ritchie, UK

One aspect of using these photoconductors as picosecond detectors that isn't discussed very often is their noise. This, of course, could be quite important if they are going to be used in high-sensitivity receivers at the end of a fibre link. It seems to me that there are one or two quite serious problems here in that the devices tend to be very small, which means that the dark currents could be quite high. There are often great difficulties in making contacts to the damaged material which could give rise to a lot of shot noise, as well as the Johnson noise you get by the inherently resistive property of the material. Have you got any comments on that?

Author's Reply

I agree with you that it is a problem to have a high sensitivity with these devices; for this reason I only presented the switching capability of this device. High speed is the advantage of this device and not sensitivity. But in principle, if you have a good dimensioning of the device, it could be increased because you could use gate width in the order of some micrometres and by focussing very tight. So you could obtain responsivities which are maybe in the order of sensitivity of PIN photodiodes, yes? I agree with you that there is a problem with this, but we did not investigate the noise properties of these devices because we are working with high optical power densities and not with optical communication systems up till now. The noise properties of these high-resistivity semiconductor materials have not been studied in detail. However, one would expect that the dark current is higher and the quantum efficiency lower (due to surface recombination processes) than in PIN photodiodes. The advantage of these photoconductor materials is their high speed; by a careful control of the electronic transport properties (e.g. by radiation damage dose) one could keep the important parameters, such as lifetime and mobility of the carriers, within a regime of operation (speed, responsivity) appropriate for its use in a high-speed optical communication system.

A.Miller, UK

I wonder what the typical length of your strip line is and if you've made any measurements of the pulse broadening due to dispersion in the strip lines?

Author's Reply

The dimensioning which we used was in the order of 10 millimetres of the transmission line, but maybe 5 mm of transmission up to the connection of the microstrip to co-ax. At this length we did not observe pulse broadening effects due to dispersion of the transmission. But you have to take into account that there can be dispersion by travelling of a short electrical pulse on the transmission line, yes? This is true. However, dispersion of ultrashort electrical pulses propagating along microstrip transmission lines have been predicted theoretically (Li, Arjavalingam, Dienes, Whinnery, IEEE MT, **30**, 1982, pp. 1270—1273) and in the meantime, also investigated experimentally. (See Proc. SPIEE Meeting on "Picosecond Optoelectronics," San Diego, 1983).

The prospects of the digital optical computer

H. Bartelt, A.W. Lohmann, J. Weigelt

Physikalisches Institut der Universität
Erwin-Rommel-Straße 1, 8520 Erlangen
Federal Republic of Germany

Abstract:

Electrons are well suited for switching operations as they occur in a logic processor. But electrons are not so easy to guide from one point to another point. On that score, photons are more suitable. Recently, photons became amenable also for logic processing. Hence, the stage is set for the development of an all-optical digital computer.

A digital optical computer has as its assets: parallel processing; global interconnections, that are favourable for architecture and algorithms. But the digital optical computer is a late-comer. Nevertheless, the optical computer has a chance to supplement the electronic computer as a special purpose parallel processor.

1. Introduction

Probably hundreds of billion Dollars have been invested world-wide for the development and for the improvement of the digital electronic computer. This development was certainly highly successful. Hence, it may be surprising that several teams around the world try to develop now an digital optical computer. What is wrong with electronics? What is so good about photons?

In chapter 2 we will compare the suitability of electrons and photons for the purpose of computing. Based on these fundamental features, we will point out in chapter 3, why photons are in some ways better suited for parallel processors. Parallel processors are needed to satisfy the ever-increasing demand for speed of computation. In chapter 4 we will briefly survey the status of digital optical computing experiments, including our own projects. In the final chapter we will express our belief that a digital optical computer has a good chance to supplement the capabilities of the digital electronic computer.

2. Electrons and photons as carriers of information

The interaction between two electrons is strong, whereas two photons, normally, do not interact at all. From this comparison follows: electronic signals are easy to switch, which is necessary to perform logic operations. But it is difficult to influence a beam of photons. We conclude: electrons are more suitable than photons as carriers of information, if logic operations are to be performed.

Logic operations are not everything that happens within a computer. A job of equal importance is the communication of signals from one point to another. When electrons travel from here to there, they have to be guided carefully by means of wires. The various wires have to be kept apart. Otherwise, electrons on different wires might interact and thereby disturb the transportation of signals. Apparently, the strong interactions among electrons was good for switching, but it is bad for communicating. For photons, it is just the opposite. Beams of photons can travel in free space, without any guiding structure, without disturbing each other.

At this point of our discussion, the electrons deserve a grade A for switching, the photons grade F. In communication, the photons are graded A and the electrons B or C. With grade F ("failure") in switching the all-optical computer seems to be ruled out. What is the fundamental reason for the grade F, or in other words, for the failure of two photon beams to interact? When two entities (electrons or photons) interact, they have to satisfy the laws of conservation (energy, momentum, angular momentum). Photons, in addition, are obliged to move at the speed of light. In other words, during the interaction of two photons one more equation has to be satisfied. The system of equations is overdetermined. Nothing happens. That situation changes, however, if a piece of exotic matter is around, while two photons meet. The piece of matter acts like a midwife, taking up the remainders of the various balances of conservation. To find the most suitable types of exotic matter is presently a very popular goal. Significant progress has been made already.

3. Electrons and photons in parallel processors

Already John von Neumann realised, that a parallel processor would be superior to a serial processor. However, a parallel processor would require a tremendous meshwork of wires, if electrons are used as carriers of information. To avoid such a meshwork, von Neumann made an ingeneous "emergency invention", which is now called "von Neumann's bottleneck"; certainly an unfair expression, because without that bottleneck, it would have been difficult for the electronic computer to get off the ground. The unfairness inherent in the expression "bottleneck" is even more apparent, if one remembers, that von Neumann also had the idea to use electromagnetic radiation instead of electrons as information carriers.

Meanwhile there exist a few quite powerful electronic parallel computers. Their potential, however, is limited in comparison with the (not-yet-existing) optical parallel computers for fundamental reasons. An electronic parallel computer may consist of an array of microprocessors, arranged on a cartesian grid. Every microprocessor is connected with its four nearest neighbours. A computer architect would prefer it, if every microprocessor had a direct line to every other microprocessor. That would require (N-1)! communication channels in the case of N microprocessors. The topology of such a network of wires would be quite complex if N is reasonably large, say logN = 20, which is desirable for image processing.

With light beams such communication networks are not unrealistic. Image processing by means of a good lens, supported by a smart hologram, does the job of simultaneous transport of one million input data (object points) to one million output locations (image points). The input data from a specific object point will appear not only in the immediate neighbourhood of the associate image point, but anywhere in the image plane, if the hologram was designed accordingly. In the language of computer architects, we may conclude: electronic parallel computers are capable only of local connections, while an optical parallel computer is well suited for implementing global connections. In other words, an optical parallel computer may be quite useful, if the algorithm of the computational problem calls for global connections.

4. State of the art in digital optical computing

In this chapter we can be very brief due to the July 1984 special issue on "Optical Computing" of the Proceedings I.E.E.E. A.A. Sawchuk and T.C. Strand /1/ present a survey and they report on their own laboratory experiments. A. Huang /2/ points out, why optical parallel computers are promissing from an architects point of view. Ichioka and Tanida /3/ present their optical logic processor, which is surprisingly simple in concept. The rest of the papers in the I.E.E.E. issue are no less important, dealing with various aspects of devices and algorithms. Our own experiments in digital optical computing make use of diffraction and scattering /4,5/.

We may conclude that the architectural aspects of the digital optical computer are well understood. Several laboratory experiments serve to explore the systems aspects. Research on the optical digital components is quite lively.

5. The chances of optical digital computing

It is too early to make a definitive statement on the chances of digital optical computing. Nevertheless, a few encouraging arguments can be made. The need for solving partial differential equations and other problems with intrinsically parallel data is steadily increasing, as can be seen by reading the January-1984-issue of Proc. I.E.E.E. on super computers. If more high speed computing power were available, more problems could be tackled.

The need for more computing power does not necessarly call for an optical computer. However, the architectural advantages of the optical computer justify an attempt to develop special purpose processors, where signals are carried by photons.

Such a development would not occur, probably, if the costs for the optical approach would be nearly as high as they were in electronics. However, the development of an optical computer does not start from scratch. For example, an ordinary burning glass, if labelled as "photonic fan-in device", is ready to be put into operation right now. A good lens combines rays in its focus with an accuracy of synchronisation as good as a femtosecond.

Beyond classical components (lenses, prisms, polarisers, beam-splitters) there are holograms, fibers, integrated-optics devices, that exist already. Also the IC technology for producing arrays of optical micro-devices can be called upon. If GaAs should turn out to be the crucial material for optical micro-devices, that would be just fine, since that material is now developped anyway for speeding up electronic components.

In summary, then, an optical digital computer would be based on capabilities, that exist already more or less in the electronic and optical industry. Such a digital optical computer could supplement the capabilities of electronic computers in solving large problems, that are parallel in nature.

One of the authors (AWL) wishes to acknowledge many stimulating discussions with Alan Huang.

References

/1/ A.A. Sawchuk, T.C. Strand: "Digital Optical Computing", Proc. IEEE 72, 758 (1984)
/2/ A. Huang: "Architectural considerations involved in the design of an optical digital computer", Proc. IEEE 72, 780 (1984)
/3/ Y. Ichioka, J. Tanida: "Optical parallel logic gates using a shadow-casting system for optical digital computing", Proc. IEEE 72, 787 (1984)
/4/ H. Bartelt, A.W. Lohmann: "Hybrid logic processing in parallel", Appl. Opt. 22, 2519 (1983)
/5/ H. Bartelt, A.W. Lohmann, E.E. Sicre: "Optical logic processing in parallel with theta modulation", J. Opt. Soc. Am. A1, No. 9 (1984);
A.W. Lohmann, J. Weigelt: "Optical logic processing based on scattering"

DISCUSSION

J.W.Haus, US

I have two questions. How much parallelism, how many parallel channels can you handle on an electronic computer? What's the upper limit? The second question is, Goodman says he doesn't believe in this digital optical computer either at least according to his talk in Anaheim, but he says that the optics can contribute interconnects between modules, or inside a chip itself, or give global interconnection in an electronic device.

Author's Reply

You asked two questions: one is the degree of parallelism in electronics and in optics; the other question was: should there be a purely optical computer or would optics take over only part of the job? In terms of parallelism you can take, as a rule of thumb, that a lens for $1000 is good for handling 1 million channels separately. If you are willing to pay more, if you are buying one of those lenses as they are used for integrated circuit manufacturing then you might have 10^7 independent channels in parallel in a lens system. So that's about the degree of parallelism that one lens can handle. Lenses may be an important part of such a system. How much you can do in electronics I don't know. I only know that NASA wants a systems of 1024 by 1024 parallel processors. (See Fischer and Schaefer in **IEEE Spectrum**, 1982). NASA would buy it whether it is optics or electronics; the main thing is to keep up with the many, many data that are coming down from space. You know the number of data coming down from satellites is of the order of 10^7 Bibles per second. So, the amount of data which is coming down from satellites is just so tremendously large that you need very fast parallel processors, and a parallel processor is appropriate because semantically the information in satellite pictures is parallel. So therefore it's quite plausible that a parallel processor can best handle parallel information which is semantically parallel.

O.K. That was one question. The other question was what Goodman said at Anaheim. Well, that's already a while ago; I don't know what he said recently. You should read his latest paper which appeared in the July isssue of **Proceedings IEEE**. Maybe I shouldn't talk for Joe Goodman. At least — yes, it is conceivable that a computer would be hybrid, that certain communication steps like perfect shuffle are done optically and then there are detectors that change the signals into electrons and after they have been processed locally by doing some logic then they are emitted again by light sources and so on. This is quite conceivable and actually in some of these circuits that are conceived and worked on, for example, at Bell the difference between photons and electrons in the material is really not so apparent any more. Or, for example, I saw devices developed at MITI in Japan and at Hitachi; there again it wasn't quite apparent any more the difference between optics and electronics. For example one device consisted of a unit which consisted of eight little photo detectors; they'd been connected in parallel in order to build up enough voltage to drive a little transistor all in the same substrate. And the output of *the transistors again triggered a light emitting* diode or laser, so this was really one unit. They are working on making whole arrays of this unit, so you might say that it's optics as electronics but in any case such a layer of parallel devices would be desirable. Apparently people are working on it in various places. Maybe in Edinburgh the ideas are different about those optical logical layers.

S.D.Smith, UK

I think you rightly commented that the global interconnection possibility of optics perhaps will turn out to be the most important idea in digital optical computing, but can you give us on example? Let's suppose that we can have a piece of nonlinear digital optical logic in one element that could address say 10^3 elements or 10^6 elements; this is now, in principle, possible from experiments. Can you give us an example of a problem where that could be useful?

Author's Reply

In Fourier transforms the output of one particular logic layer doesn't go through many steps so you don't have a fan-out ratio of one thousand. I don't know of such a problem. But from what I hear about structural analysis in weather forecasting, there, sometimes, the fan-out ratio in terms of the algorithm can be very high. This might be a problem for which an optical parallel computer is suitable. Furthermore, this might be a suitable first project, because such a weather forecast computer would be a somewhat specialized computer, which is all right because a weather forecast station does weather forecasting day after day, so they can afford to have a somewhat specialized computer. Furthermore when people say, well, but your optical computer will be a special-purpose computer, it doesn't really worry me so much because even electronic computers are to some degree all special purpose computers. You wouldn't take the computer out of a bank office and then use it for scientific computations. The configuration of computers is always at least in partial degree special purpose, i.e. the term 'special purpose' for me is not a dirty word.

Unidentified Speaker

I think that in several applications, such as in my personal field of radar signal processing and sonar signal processing and such things, you have a huge amount of data coming in in parallel. Nowadays we just pass them through one bottleneck. It's a good idea to have a parallel processor which is capable of performing matrix algebra, etc., but it must be programmable and must have the accuracy of a digital computer. We are not satisfied by the existing logic computing devices.

D.Bosman, Ne

In your presentation you have kept mainly to the structure of a computer and not to the metrical information, the amplitude or magnitude. We may not assume that all data paths have an equal length. So, we must also assume that the inputs vary in magnitude. What kind of dynamic range do you foresee, which is acceptable for the optical computer?

Author's Reply

Well, the advantage of binary computation is, of course, something which an optical computer should try to make use of. And when people talk about digital optical computers they usually think of binary signals. Some people do not; there are some acousto-optical psuedo-binary computers conceived, and in **Fortune Magazine**, you could probably read a big article on this. These might be very special purpose computers — I don't know — that's in the long range. The future of optical optical computing is questioned by you in saying that if you have data coming in in parallel some data are upheld because they have to undergo some processing, but some other data are synchronous at first — they are not delayed by processing. So we have a problem of synchronization of large data fields; for that purpose one needs tapped delay lines. Optical tapped delay lines with parallel features, even controllable, is already one of the subjects we tried to address; at least we have some ideas how this can be done. It's definitely the truth that such a device would be necessary in order to synchronize data streams.

D.Bosman, Ne

Isn't it true that in an optical computer you would need a new device which is not necessary in the electronic computer like a magnitude regeneration unit?

Author's Reply

Well, in an electronic computer you also need a refresh stage to bring up the signal to a certain defined level. So what we definitely need is something like a non-linear hard clipper which accomplishes this purpose. So I think in the set of schemes we call indirect linear logic, we would need really hard limiters and at the same time, converters from incoherent to coherent light. If I do have a layer which converts incoherent to coherent light and at the same time does some hard limiting, bringing it all up to the same level, then I think I can set up a total system.

J.P.Dakin, UK

In the description of your computer architecture you said that you would have a random shuffle arrangement using holographic type elements in the system. Either of these could be some sort of fixed elements which would deflect perhaps in a given direction and then you don't have the freedom for random access or they may be perhaps addressable and programmable. You may need a digital electronic computer with the same capacity of the computer you're trying to design in order to be able to randomly address such an array of holographic elements. Could you perhaps go into details on the problem of this, which seems to be perhaps the main stumbling block in the optical computer, as far as I can see.

Author's Reply

Yes that's right. The computer should be programmable. So the first step is the opportunity to simply bypass the perfect shuffle. This can be done by having, for example, a Wollaston prism which either swings the whole beam to the left and it goes through the perfect shuffle, or to the right, and then it simply bypasses the perfect shuffle. The next thing is that you want to apply the perfect shuffle, not to the whole field, but you want to apply it only to one-half field and separately to another half field. Or maybe only to one quarter field and so on. If you do this sub-division of perfect shuffling into partial fields — binary partial fields — then you come up with something like bunyon structures and benish codes and things like that. So it is possible to generalize it so the basic component that does the control is, for example, a Wollaston prism which can cause either individual beams or whole bunches of beams to swing to the left or to the right and then find the proper hologram or whatever unit might be that does the shuffling either for the whole array or only for partial rays. That provides flexibility. In addition it's so that even if you do always the perfect shuffle with the whole field but in between you always do some nearest neighbour operations where you have either exchange of nearest neighbours or let them just go straight or possibly every second one is eliminated or something like this, then this also gives a large amount of freedom in rearranging data structures or readdressing.

E.Spitz, Fr

My question is about this perfect shuffle, too. You were saying in an electronic matrix for example, if you want to propagate a signal from the left to the right, it goes through several cycles. This is true because you have got elementary two-position switches, and you want to have complete flexibility and so you go through all these steps. You could (in electronics) put in parallel to each layer derived value and then you would go straight, but you don't want to do this because this makes the preparation too difficult. Now if I'm coming to your perfect shuffle and if again you want to be flexible, you will need some elements and you were speaking about the Wollaston prism which will have only two possibilities, probably. So you will be in the same situation because either you will put all the things in parallel and you will have only the path for one ray, or you will cycle it. Then you are in the same situation as an electron system. So I understand the general idea, but I have the impression that for the same flexibility you will be in the same situation in optics as in electronics. Am I right?

Author's Reply

I hope not. I think what you are implicitly talking about has something to do with the so-called single instruction multiple data-stream parallel computers and multiple-data-stream computers. If you uniformly apply all the same rules, the so-called single-instruction multiple-data-stream system is relatively easy to implement and that one would do first. The applicability of single instruction multiple-data stream is probably good enough for the beginning because if you are investigating images and you are looking for a certain feature, then you are asking the same question for the upper left corner and the lower right corner of the image field, so the single instruction multiple-data-stream approach for the time being is at least the first goal that is useful enough

E.Spitz, *Fr*

I understand this but that is true in electronics, too; you've got cable logics which are less flexible and then it goes much more rapidly. For example the fast Fourier transform, where you have cabled logics which are much faster. So you are exchanging flexibility for some more specialized calculations. I think, if you want to be honest, we have to always compare the same type of flexibility and same type of calculability.

DEFORMABLE MIRROR NEAREST NEIGHBOR OPTICAL COMPUTER

A. D. McAulay, PhD
Texas Instruments Inc., Central Research Lab., MS 238
P.O. Box 226015, Dallas, TX 75266

SUMMARY

Many problems are efficiently computed on nearest neighbor machines because they involve regular grids or grids which may be envisaged as distortions from regular. Examples include image processing and finite differences and finite elements which dominate engineering and scientific computation. Last year I proposed an optical nearest neighbor machine concept with 5 bit accuracy, (1). The simple design in this paper uses residue number arithmetic and two deformable mirror arrays of 1000 by 1000 elements to achieve 15 million operations per second with 32 bit accuracy. I assume that an array may be set in 8 mseconds. Duplication of equipment by 64 times enables one billion operations per second because perfect parallelism is achievable with residue numbers and nearest neighbor concepts. It is anticipated that the increased use of deformable mirror devices in computer peripherals will advance the technology to the point where designs such as that proposed will be practical. The proposed scheme appears to be superior to outer product and optical matrix schemes for nearest neighbor computation.

INTRODUCTION

Finite elements and finite differences represent the dominant computational cost in numerical computation for science and engineering. Many of these problems involve regular grids or grids which are distorted from regular. Such problems, along with many image processing problems, may be efficiently mapped on nearest neighbor connected machines, figure 1. At the same time, many of these problems cannot be tackled today because present day machines are too costly or too slow. In other cases the power requirements or the weight are too great. Examples of problems requiring faster computing arise in inverse problems such as: the determination of the source associated with a field, the determination of the media through which a field propagates (as in geophysics) or the determination of stiffnesses in the optimization of a structural design. These problems require many iterations of the normal finite element solutions and consequently, only scaled down versions of the real problem are possible today. An interesting example is the updating of structural and aerodynamic models on board an aircraft so that the pilot may use the models to predict the effects of certain operations. This could represent a useful aid during test flights or in the event of a structurally damaged aircraft.

Previously, I hypothesized an optical machine, (1), using two dimensional spatial light modulators in an analog mode which provided an accuracy of 5 bits. The configuration proposed here is likely to be feasible sooner because it is simpler and more useful because it has greater dynamic range. It runs nearest neighbor type computations efficiently and thus takes advantage of the regularity arising regular grids. I consider computations first to finish and then to not to overlook bottlenecks which might arise in input-output or elsewhere in the computation.

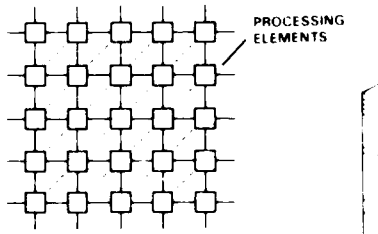


Fig. 1 Nearest neighbor array

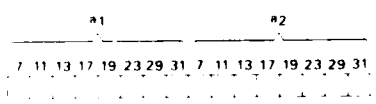


Fig. 3 Residue number coded vector

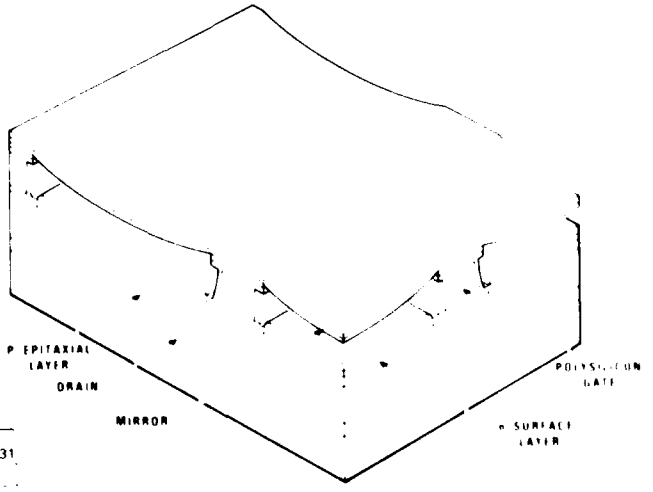


Fig. 2 Deformable mirror pixel

Spatial light modulators have been investigated by many people, [6]. The proposed processor uses spatial light modulators of the deformable mirror device type, DMD's. Examples are membrane light modulators which have been constructed in 1-D, [6,7], and 2-D arrays [8,9] and cantilever beams [10]. Membrane light modulators consist of a semiconductor device with an array of 128 by 128 or more transistors, figure 2. A reflecting membrane covers the surface of the array. Activation of a transistor causes the membrane to be locally attracted and results in a dip above the transistor. The depth of the dip is dependent on the voltage applied to the transistor. The transistor array is set, a row at a time, by applying appropriate voltages to drain terminals across the top of the device and clocking the appropriate row. A cantilever beam device is similar except that above each transistor it has small plates attached by one corner. Energising the transistor causes the corresponding plate to bend down.

DMD's are especially targeted to peripheral devices such as displays. This could be used to advantage in iterative schemes where it may be desirable to see the results as the computation progresses. There are significant benefits in reducing the number of different technologies used in a system. Progress in optically addressable 2-D arrays could lead to a system in which input is applied to a DMD with a light pen via an optical system. Output is directly from a DMD to a screen and computations are performed at high speed with DMD's. Even hard copy may be performed directly be addressing the DMD head of a printer. At this stage the system could be very competitive with conventional electronics for which many interfaces and different technologies are required.

The proposed processor uses residue number system theory [11], RNS, to achieve dynamic range. RNS has previously been investigated for application in optics, [12,13,14]. A convertor is used during iterations to return the outputs of the residue number computations back to the range of their moduli. An analog design seems simplest because the requirements for accuracy are reasonable with the limited range of the numbers involved in nearest neighbor residual number computations. Optical methods are proposed in reference 13 but electronics is required to convert from intensity to displacement and back.

Other researchers, such as P. Guilfoyle [15,16], plan to achieve the desired performance with an acousto-optic device used in a systolic mode [17,18] with digital multiplication by analog convolution to achieve sufficient dynamic range, [19]. Acousto-optic cells are used widely because they are well developed and exceptionally well suited to spectral analysis applications in which data is applied directly, or with frequency shift, to the acoustic transducer. In addition there has been a trend to propose sytems to take advantage of more readily available 1-D light modulator technology [20,21]. Many of these systems do not appear to be too practical. The outer product concept for multiplication does not take advantage of the nearest neighbour nature of the computation and because outer products are accumulated on a time integrating device, the dynamic range of this device limits the size of matrices that may be handled. Others have investigated the solution of partial differential equations with analog processors [22] and with systolic arrays [23].

CONCEPTS

System overview

A host computer with a terminal are shown at the top of figure 4. I observe that humans have a limited capacity for input and output, so that as the problem size and complexity increase I do not expect the input output to increase in proportion. I assume that the setting up of the matrices and the output operations are not a significant part of the computation time for large finite element problems. These operations may then be efficiently accomplished in the host computer. The matrix elements are, for example, functions of the stiffness of a structure or the propagation medium properties.

In the case of a regular grid, or a grid distorted from regular, a nearest neighbor computation is suitable. This requires that every point in a matrix, called the field matrix, may be updated by adding specified amounts of values at one or more of the neighboring points. The specified amounts to be added from a neighbor to the north of each point generates a 'north' coefficient matrix. This represents a spatially varying filter, as the amount to be added to each point varies according to the position of the point. Similarly other coefficient matrices may be considered for east, west and south directions. In the case of finite elements, values at a point represent a field over a finite element region, consequently, additional matrices are required. In the use of triangular elements a north-east and south-west matrix will suffice.

An initial estimate of the field, e.g stress or magnetic field, is also required for iterative schemes. In cases such as that arising in geophysical prestack inversion, [3], where a region to be estimated overlaps that estimated earlier, a good estimate is known and the iterative technique is preferred. In cases, such as those frequently arising in structures, where the problem is to be repeated many times with different loads, LU decomposition is more numerically efficient because the decomposition need be performed only once. Optical

methods of computing such numerical techniques have been proposed [11]. Some modification of the proposed design will be required to compute Gauss Elimination, LU Decomposition and non regular grid problems efficiently.

The box at the lower left in figure 4 uses parallel 2-D element by element optical multiplication, [1], to perform the main calculation. The processor and its use for performing nearest neighbor computations are described later. The ouput intensity measured by the CCD is modulod back to an intensity within the range of the residual number system in the lower right box. This is explained in the sections on the residue numbers and the modulo operation.

The new intensity is also fed to the host for convergence evaluation and display purposes. Convergence evaluation is fast because it does not require transformation back to decimal numbers and may be performed in parallel. Each residue number may be compared with its value on a previous iteration separately. The convergence evaluation is available to the user for control of termination if he so desires. The field results are displayed at infrequent intervals and it is anticipated that the user will wish to examine these interactively at the end of the computation. The convergence evaluation and intermediate displays may be performed at the same time as the optical part of the processor is continuing to compute further iterations. The conversion time is small relative to the computation time because large problems requiring the speed of this processor will perform hundreds or thousands of optical iterations before the results are of interest. Consequently, the speed of the convergence evaluation computation and the conversion to decimal for display are less critical and may be performed in the host.

Residue numbers

The dynamic range of the DMD's is of the order of 5 bits which is insufficient for partial differential equation solution. Two methods appear most promising for increasing the dynamic range, digital multiplication by analog convolution in a binary convolver and the use of residue number systems. The reasons why I selected the latter are presented in this section.

Residue number systems have been built in digital electronics at Lockheed, Martin-Marietta and Mitre [11] and optical systems have been proposed [12,13, 14]. The principle advantage is that once numbers have been converted to a code, coded numbers can be added and multiplied by separate operations on each element of the code and no carries between elements of the code are necessary. Let us perform operations 7 + 3 and 7 x 3 for the purpose of illustration. Select a set of relatively prime numbers such as (3,4,5) for a moduli set. The residue code for a number is determined by modulating with each element of the

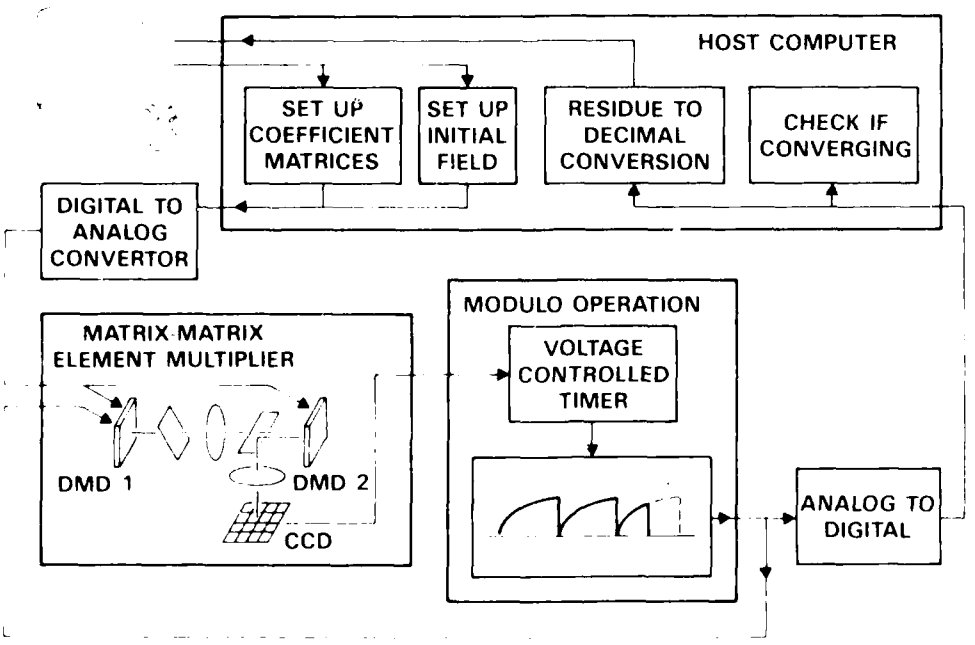


Fig. 4 Deformable mirror nearest neighbor optical computer

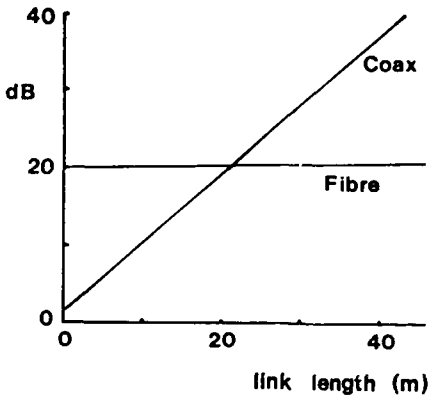


Fig. 8 Typical Insertion Loss of R.F. Links (5 GHz, with fibre attenuation 1 dB/km and coaxial cable attenuation 1000 dB/km)

It has been calculated[23] that, using optimised values of $\mathcal{R}$, K_L, R_L, and r, the intrinsic insertion loss of the electrical-optical-electrical conversion could be reduced to a minimum of 10 dB.

The overall S/N for a single optical fibre link[22] depends on the photodetector shot and thermal noise together with thermal noise associated with the following amplifier, and on the laser a.m. noise (laser intensity noise). For signal transmission at GHz frequencies, laser noise is generally dominant for short links which have relatively high optical power at the detector. When the detected optical power is low, detector and amplifier noise usually dominates (Fig. 9).

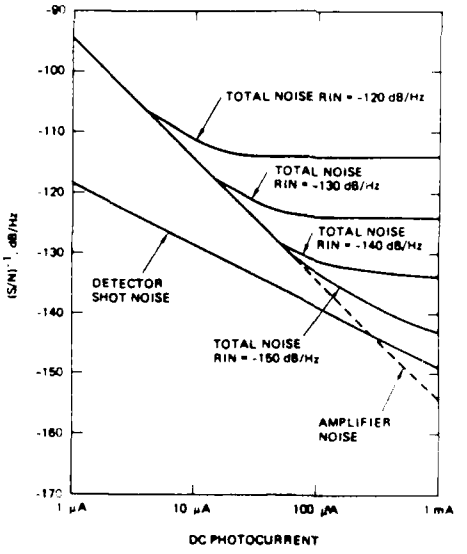


Fig. 9 Typical S/N of a Fibre Optic R.F. Link (Ref. 22). RIN represents the level of laser relative intensity noise.

The conclusion as regards fibre optic signal distribution manifolds is therefore that the required S/N ratios for radar r.f. signal distribution in a phased array may be obtained, but that the number of output ports is severely limited, typically to about 10, from a single laser of the usual 10 mW output power. More output branching may be provided if a degradation in S/N is allowed (Fig. 10).

A principal advantage of optical fibres lies in their weak sensitivity to temperature. Changes in temperature of a fibre affect the signal propagation delay through a thermal expansion in length, L, and a change in refractive index, n. For a length L of fibre, the change with temperature T in insertion phase, ϕ, for a modulation frequency, f_m, of the optical carrier is:

$$\frac{d\phi}{dT} = \frac{2\pi L f_m}{c} \left(n\alpha + \frac{dn}{dT} \right)$$

where c is the velocity of light in vacuum and α is the fibre thermal expansion coefficient. Both α and dn/dt typically have values around $10^{-6}/^{0}C$ and thus for L = 10 m and f_m = 3 GHz, $d\phi/dT \approx 0.1^{0}/^{0}C$. This is approximately an order of magnitude better than for coaxial cable[20], microstrip or stripline structures.

Surprisingly, one of the most difficult problems encountered with fibre optic networks is still that of splitting and combining optical paths. Relatively low loss (eg. 1 dB) single mode connections may be made[21], but branching networks (eg. star couplers) still are difficult to manufacture. Again, the solution may only lie in integrated optical analogues of conventional waveguide couplers.

3. R.F. SIGNAL DISTRIBUTION

3.1 Outline Considerations for a Phased Array Signal Distribution Manifold

The likely requirement of future active phased array systems, as identified earlier, is for distribution of r.f. signals at a level of a few mW. This can be done in any of the conventional transmission line media such as waveguide, coaxial cable, stripline or microstrip. Disadvantages are the relatively complex mechanical construction involved, the significant mass, and the poor stability of insertion phase with temperature. Loss can be significant, but since path lengths are relatively short, this does not usually exceed more than a few dB. The trade-offs to be made in considering whether a fibre optic manifold is a viable replacement for a conventional manifold lie chiefly in considerations of loss and stability, together with the fact that the fibre optic network can carry other signals as well.

An approximate power budget for the optical signal distribution manifold (Fig. 7) may be calculated easily and is given by[16]:

$$\frac{P_{out}}{P_{in}} = \left[\frac{\mathcal{R} \alpha_c \alpha_t \alpha_b K_L}{N} \right]^2 \frac{R_L}{r}$$

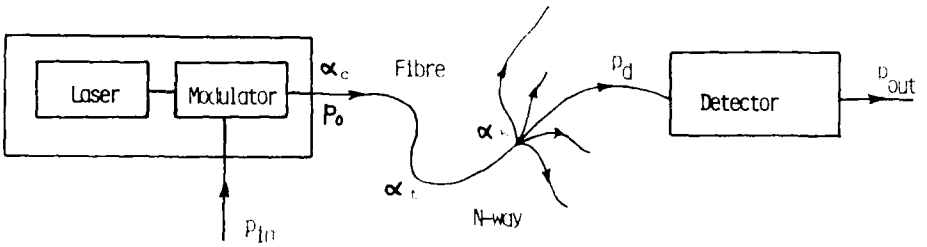


Fig. 7 Fibre Optic Manifold for R.F. Signal Distribution

where P_{in} is the r.f. signal input power, P_{out} is the r.f. signal output power; α_c, α_t and α_b are optical transmission coefficients representing losses in coupling (laser-fibre, fibre-detector), attenuation in the fibre, and in fibre branching nodes, respectively. N is the number of array elements to be fed, R_L is the detector load resistance and r is the laser dynamic slope resistance.

For typical values such as $\mathcal{R}$ = 0.25 A/W, K_L = 0.2 W/A, R_L = 50 Ω, r = 4 Ω, α_c = 0.5, α_t = 1, α_b = 0.7:

$$\frac{P_{out}}{P_{in}} = -24 - 20 \log N \quad (dB)$$

For a single, unbranched optical link, the insertion loss of 24 dB is fairly typical of what has been achieved in practice[22,23]. It can therefore be seen that a severe initial penalty is paid as a result of the electrical-optical-electrical conversion, but that the overall link loss in long optical links can be lower than for conventional links (Fig. 8). It should also be noted that the signal power on the network branches decreases inversely as the square of the number of output ports because of the square law relation between optical intensity and signal current.

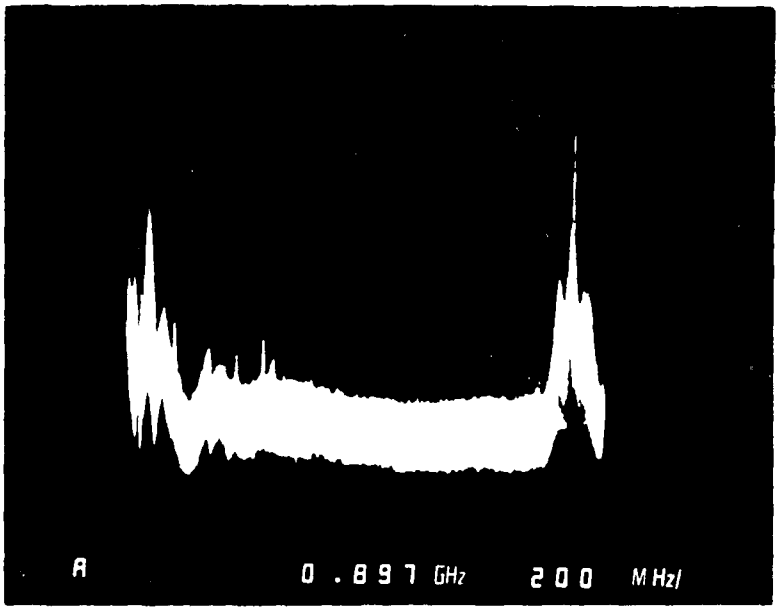


Fig. 6 Detected Output of Two Laser Beams Phase Locked to a Microwave Difference Frequency close to 2 GHz

2.3 Detectors

The most common optical detector is the pin diode, which provides an output photocurrent, I_p, approximately a linear function of the incident optical power P_d[17]:

$$I_p = \mathcal{R} P_d$$

where $\mathcal{R}$ is the responsivity, typically of value 0.4 A/W.

Noise in the detector arises from dark current, shot and thermal effects. The minimum optical power level $P_{d,min}$ at which a given S/N for a signal of bandwidth B may be obtained, assuming thermal noise dominates, is:

$$P_{d,min} = \frac{2}{\mathcal{R}} \left[\frac{4kTB}{R_L} \left(\frac{S}{N} \right) \right]^{\frac{1}{2}}$$

with R_L the detector load resistance, k Boltzmann's constant and T the noise temperature of the detector (1000-3000 K, typically).

Detector frequency response is limited by RC and transit time effects in the diode; this tends to imply the use of low values of R_L. Recent developments using GaAs Schottky barrier structures[18] have shown that devices with corner frequencies of at least 20 GHz and maybe in excess of 60 GHz can be realised.

Avalanche pn junction detectors give higher sensitivity through their internal gain; FET's also can act as photodetectors with gain[19]; however, it appears that optimum performance is obtained with a pin detector followed by a low noise MESFET. Some performance optimisation may be carried out by careful impedance matching of the MESFET input as the diode load.

2.4 Optical Fibres

The use of single mode or multi-mode optical fibres in a signal transmission link depends very much on the nature of the signal to be carried. The speckle pattern in a multi-mode fibre is very sensitive to vibration or mechanical distortion of the fibre and this can significantly degrade the close-to-carrier S/N.

To preserve high S/N, a single mode fibre is required and it may also be necessary to use polarisation filters to ensure polarisation purity. Single mode fibres also avoid problems of modal dispersion of the signal, the dominant dispersion effects being due to the fibre material alone. Such dispersion effects may usually be totally neglected for radar applications where likely transmission path lengths for signals are relatively short. Typical (bandwidth)(length) products for single mode fibres are in excess of 10 GHz.km for the most likely optical wavelengths used. The optical signal attenuation may also usually be neglected since it does not exceed a few dB/km.

with K_L a constant, typically of value 0.2W/A. In general, a mixture of a.m. and f.m. is produced due to the effect of the bias current on the light intensity and on the temperature and refractive index of the laser cavity[14]. Modulation bandwidths up to 6GHz have been achieved from commercially-available lasers, and there is evidence that new laser structures may allow modulation bandwidths in excess of 10 GHz.

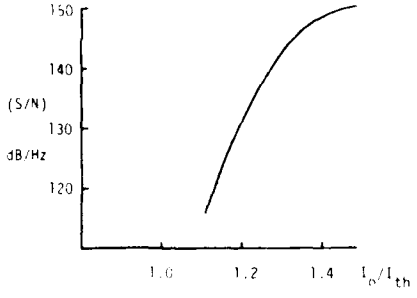


Fig. 4(a) Laser A.M. Noise (Relative Intensity Noise) as a Function of D.C. Bias Current Normalised to Threshold Current

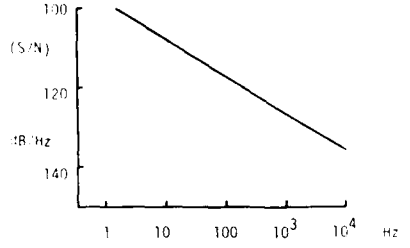


Fig. 4 (b) Laser A.M. Noise (Relative Intensity Noise) as a Function of Frequency off Carrier (Ref. 12).

External modulators offer another means of modulating the laser output. Integrated optic versions[15] have demonstrated modulation bandwidths up to about 20 GHz, but for relatively low optical power such as 1 mW, beyond which optical damage effects occur.

Thus, it may be said that r.f. modulation of optical cariers at optical power levels of some mW is practicable only at up to about 6 GHz at present. Radically new developments will be necessary to extend these capabilities by significant margins. One possibility[16] is the use of two c.w. semiconductor lasers phase locked by a microwave frequency offset in a heterodyne loop (Fig. 5). Experiments have shown the viability of this scheme (Fig. 6), but the use of any technique involving coherence at the optical frequency still brings considerable practical difficulties which may only be solved by integrated optics.

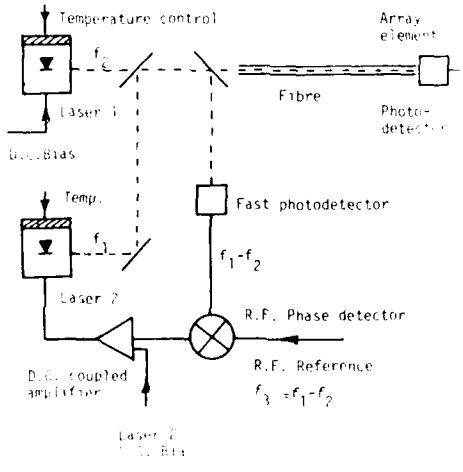


Fig. 5 Heterodyne Phase Locked Loop With Two Lasers for Microwave Modulation of Optical Carriers

The i.f. beamformer might thus operate at a frequency in excess of 100 MHz with up to 50% bandwidth. Both the transmitter and receiver systems in satellite communications work with relatively constant signal levels and thus dynamic range requirements are very modest, rarely exceeding 30 dB.

2. SIGNAL DISTRIBUTION USING OPTICS

2.1 Advantages and Disadvantages

Active phased arrays are likely to require signal interfaces at r.f., i.f., and baseband. These interfaces will be made from a central control unit to each active module in the array aperture. At one time these modules were the main cost component in the overall system due to their complex and precision tolerance construction. The advent of monolithic microwave integrated circuits and VLSI chips has, however, simplified these modules, making packaging, connectors and signal interfaces a relatively much larger component in overall cost and reliability considerations[10]. Considerableadvantage could be gained by implementing all signal interfaces on a single fibre optic link.

Signal distribution on optical fibres has a number of advantages:

(a) the optical carrier is at such a high frequency that signal bandwidths are a very tiny fraction of the carrier frequency; this gives excellent constancy in the transmission properties of the link over the required microwave operating frequency range.

(b) a number of signals may be multiplexed on to the optical link using different optical wavelengths (wavelength diversity multiplexing).

(c) the distribution network is flexible and of low mass, which is important for airborne, space or naval masthead applications.

(d) the distribution network has a high immunity to electromagnetic interference and cross-talk.

(e) the insertion phase variation with temperature for microwave signals on an optical carrier in a fibre link is almost an order of magnitude lower than for the same signals on a coaxial cable.

However, there are penalties to be paid with optical fibre links in terms of the loss in the electrical-optical-electrical conversion process and limitations in dynamic range that arise through these conversions. Following sections seek to clarify the limitations.

2.2 Sources

Gas lasers have narrow linewidths and provide optical output power in visible and infra-red regions at up to the watt level or more; they are, however, prone to low frequency instabilities, are bulky and require external modulators. Solid state sources are therefore preferable and the choice currently lies between light emitting diodes and lasers. The linewidth of an LED is very broad ($\sim$10 nm), its output power usually low ($<$1mW) and its modulation frequency response limited ($\sim$200 MHz for most devices). It is therefore best suited to baseband or low i.f. applications.

Semiconductor lasers are the most suitable source currently available. They now provide output powers of tens of mW at a range of wavelengths (850, 1300, 1500 nm). Multiple stripe devices have yielded output power of several hundred mW and pulse power of several W[11]. Early semiconductor lasers were multimode devices with broad spectra, but many available devices now operate in a single mode with a line width less than 1 MHz. Laser noise has been extensively studied for the communications application; the noise arises from a variety of causes, such as shot and recombination processes, a typical noise characteristic being as shown in Fig. 4[12].

Far-from-carrier S/N levels greater than 130 dB (in 1 Hz bandwidth) may be obtained and this satisfies the usual radar reference source requirements; close-to-carrier noise shows a typical 1/f character with a corner in the region of 10 kHz. However, laser noise is a strong function of optical feedback into the laser cavity from components in the optical path and may in many cases dominate the overall noise performance of an optical link. Normally, a.m. noise is the important consideration, since phase noise only becomes important in coherent systems.

Semiconductor lasers may be directly modulated through their bias current and show a dependence close to linear for the optical output power, P_0, as a function of bias current, I_0, above the threshold bias for lasing[13]. Thus:

$$P_0 = K_L I_0$$

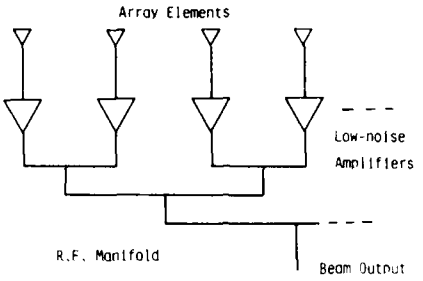


Fig. 3 R.F. Manifold in Receive Mode for Active Phased Array

manifold is more acceptable than in the transmit mode. Interest in bistatic radar systems[6] or systems where it is desirable to locate the radar signal processing equipment remote from the radiating aperture has led to interest too in relatively long signal distribution links (100 m - 10 km) for the received r.f. signal, even from conventional mechanically steered radars. An obvious requirement here is low transmission loss, less than a few dB/km, for the r.f. signals.

Many future radars will undoubtedly use i.f. manifolds for beamforming, particularly in receive mode, as is the case with the Marconi "Martello". Such beamformers make the achievement of tight phase and amplitude tolerances much easier and allow flexibility in multiple beam formation. Since amplification can easily be provided at intermediate frequencies, loss in such beamformers can be acceptable. The main constraint on the intermediate frequency is that it should be sufficiently high to pass the signal instantaneous bandwidth, which could be as high as 20 MHz, without difficulty. Typical intermediate frequencies are thus of the order of 100 MHz.

The ultimate in flexibility for the future, however, will probably lie in digital beamforming techniques[7] which only require the provision of constant level high stability reference signals to downconversion mixers prior to signal digitisation. This is covered by the consideration of r.f. signal manifolds earlier in this section.

A final requirement of array radars in terms of signal distribution is that of control signals, in almost all cases digital, to set phase and gain in the individual element or sub-array channels. The very maximum requirement here would be for a new beam direction every pulse repetition interval, unless schemes involving receive beam steering within the interval were to be implemented. For a 1000 element array and a 1 ms p.r.i., the data rate for each element is a few kb/s and for the array overall it is a few Mb/s. The exact numbers depend on the number of bits of precision in phase shift required and whether amplitude control is also needed.

1.2 Satellite Communications

Most current satellite communication systems use a single global beam to cover the earth's surface from geosynchronous orbit. This is obtained from a single radiating aperture such as a parabolic reflector or a horn. The limited channel capacity, and in the military context, lack of security, of global beams has led to the need for systems which can provide multiple beams on the earth's surface[8]. For the enormous market of communication with low gain mobile terminals on land, sea or air, future satellite systems will have to produce high gain agile beams. The satellite antenna will therefore have to use phased array concepts, either in the primary radiating aperture or in the feed for a reflector imaging antenna system. Phased array apertures are already at advanced stages of development, for example, in the DSCS III military satellite and the ESA-sponsored Multi-beam Array Model civil communications project[9].

Signal distribution requirements are similar to those in radar. Active array concepts are particularly favoured because of the limited and gradual degradation in performance as components fail. Many performance aspects, such as good inter-modulation performance, are more readily achieved by distributing lower power r.f. amplifiers over the radiating aperture rather than passive distribution of power from a single high power TWTA. Since too, communication channels are usually assembled at baseband or i.f. into groups for a particular beam, there is significant benefit in using i.f. beamforming systems.

Thus, an on-board satellite communications antenna of the future might typically require r.f. signal distribution to individual array elements at the mW power level with phase and amplitude stability of a few degrees and few tenths of a dB. Common frequencies at present are 1.5 GHz, 7/8 GHz and 12/14 GHz, with systems usually requiring only a few per cent r.f. bandwidth for the separate transmit and receive bands. However, interest in the 20/30 GHz and 44 GHz bands is increasing rapidly. The beamforming system, if of the i.f. type, would need to operate at a frequency sufficiently high to cover the transmit or receive band; because of the requirement for full duplex working, separate transmit and receive beamformers would normally be used.

In the transmit mode, the r.f. feed to each amplifier in the array aperture needs to be at a level of at least some milliwatts, otherwise the amplifier gain to achieve output power at the tens or hundreds of watts level becomes inconveniently high. To some extent, the level of the input signal is limited by the output power capability of a single reference source solid-state amplifier divided by the number of individual array element power amplifiers to be fed. (Fig. 1).

Stability of the transmit frequency is, of course, of prime importance particularly with regard to coherent processing; a typical modern radar would be expected to have a carrier signal-to-noise ratio, S/N ~ 130 dB in 1 Hz bandwidth. To achieve low sidelobes, which is currently important for lowering the probability of intercept, good control of phase and amplitude over the array aperture is necessary.

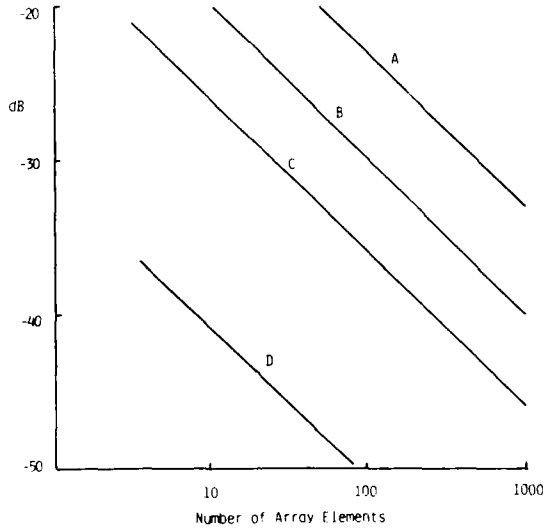


Fig. 2 RMS Sidelobe Level as a Function of Number of Array Elements

	RMS Phase Error	RMS Amplitude Error
A:	20^0	2 dB
B:	10^0	1 dB
C:	5^0	0.5 dB
D:	1^0	0.1 dB

Fig. 2 shows typical tolerances on these quantities for given achievable sidelobe levels and array dimensions.

Currently the typical radar signal r.f. tunable bandwidth is of the order of 10% to provide frequency agility, but future systems may require even greater bandwidths, possibly up to an octave.

Thus, a typical requirement for a future array r.f. distribution manifold might be to deliver to some hundreds of output ports a signal at the 10 mW level with S/N ~ 130 dB (1 Hz BW) and long term stability of 10^0 and 0.5 dB. The severity of this task is naturally also a function of the radar frequency which would most likely lie within the 1-18 GHz range at present, with a preponderance towards the 1-5 GHz band of frequencies.

In the receive mode, the r.f. manifold may be used to combine the signals from array elements or sub-arrays to a single output port, or several beam ports in the case of multiple-beam arrays. Alternatively, it may be used to distribute the local oscillator signal at constant phase to each element. It is in receive mode that high dynamic range becomes a key requirement of the signal distribution path. A typical radar receiver output before processing may have to encompass signals over a dynamic range in excess of 100 dB. In individual outputs of elements or sub-arrays in an array radar, the dynamic range requirement of each path in the manifold is reduced in proportion to the number of elements, but still represents a very large dynamic range. The phase and amplitude stability requirements of an r.f. manifold in receive mode are likely to be more severe than in transmit mode due to the need to obtain very low sidelobes and thereby reduce the susceptibility to jamming. For this, and in order to implement adaptive nulling of interference sources, as will undoubtedly be required in all future military radars, tolerances may become as tight as 1^0 and 0.1 dB. A typical future system, if it used an r.f. manifold on receive, would have low-noise amplifiers at each element or sub-array to define the system noise figure (Fig. 3). Thus a certain amount of loss in the

Neumann made an ingeneous "emergency invention", which is now called "von Neumann's bottleneck"; certainly an unfair expression, because without that bottleneck, it would have been difficult for the electronic computer to get off the ground. The unfairness inherent in the expression "bottleneck" is even more apparent, if one remembers, that von Neumann also had the idea to use electromagnetic radiation instead of electrons as information carriers.

OPTICAL TECHNIQUES FOR SIGNAL DISTRIBUTION AND CONTROL IN ADVANCED RADAR AND COMMUNICATION SYSTEMS

J. R. Forrest
Microwave Research Unit
Electronic & Electrical Engineering Dept
University College London
Torrington Place
LONDON WC1E 7JE, UK

SUMMARY

The paper outlines the signal distribution and control requirements of some typical advanced radar and satellite communication systems. These involve the ability to handle:

(a) r.f. reference signals at frequencies up to at least 20 GHz with high stability and signal-to-noise ratio.

(b) i.f. signals of large dynamic range, typically at least 70 dB, with relatively high fractional bandwidth.

(c) digital control commands at the Mb/s level.

A number of research teams have investigated exciting and elegant possibilities for the use of optical frequencies as carriers for this signal information and its processing. However, many schemes proposed have not been fully assessed in terms of realistic system requirements which place severe constraints on noise, stability and dynamic range. The paper reviews much of the current work and attempts to put possibilities into perspective, taking account of the practicalities.

Results from experiments at UCL on phased array radar signal distribution and satellite beamforming systems using optics are given to illustrate the principles of optical signal distribution and control.

1. INTRODUCTION TO SYSTEM REQUIREMENTS

1.1 Radar

Radar systems have progressed substantially in the last few years. Phased arrays, such as the Hughes "Firefinder"[1], the Marconi "Martello"[2] and the Plessey AR320[3] are in production for tracking and surveillance applications. Multiple beam formation, as used in "Martello", shows the importance of advanced signal processing in modern radar and the need to access individually a large number of parts of the array aperture (individual elements or sub-arrays). Solid-state radars such as the GE592[4], the Westinghouse 2000[5] and newer Marconi "Martello" versions reduce the need for distributing large amounts of r.f. power to the array aperture through the use of power amplifiers behind the aperture itself. Thus, it is likely that a majority of future array radars will only require relatively low level signal distribution to, and return from, the aperture. However, this signal distribution channel must allow the passage of r.f. signals with high amplitude and phase stability for the transmit and receive mode or local oscillator signal for receive downconversion, for i.f. return signals with high dynamic range if downconversion at each element occurs, and for baseband signals to control phase and amplitude at active modules in the array aperture.

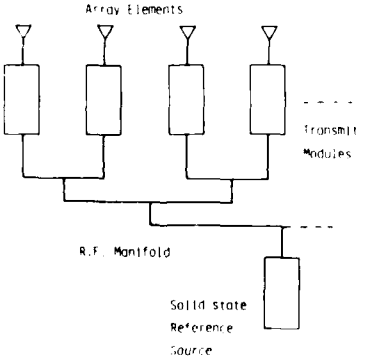


Fig. 1 R.F. Manifold for Active Phased Array

DISCUSSION

S.A.Collins, Jr., US

What's your answer to the question of residue arithmetic and division?

Author's Reply

There are several things I didn't mention. You can't do the division. It turns out the only thing you need to do in some of these problems is normalization of some sort; that can be done by computing the reciprocal and then multiplying by the stored reciprocal. You're right — it's a problem for residue arithmetic. There's another problem with the deformable mirrors and that's representing negative numbers; it arises because intensities are used for performing multiplication. If in this sort of problem it turns out that you can arrange it so that one of the things that you are multiplying by is always positive, you can do the negative products first and then do the positive products. That's messy and that's another area that needs some attention paid to it, exactly how to do it.

J.P.Dakin, UK

Could you perhaps discuss the potential problems of diffraction limitations of deformable mirrors and also the optical loss problems in going through iterations in the system, please?

Author's Reply

Well, I'm not an expert on the devices. I think on the diffraction thing you need to look at the references. I have one in the paper (also see "Advances in Optical Information Processing," by Dennis Pape and Larry Hornbeck, SPIE **388**, Jan. 1983). We feel that's manageable, that's not going to be a problem. Your other question on optical loss problems: well, if you notice in that loop I'm using some electronics so I'm not losing, I'm putting power back in in the electronics. But if I was trying to do the whole thing optically that could get to be a problem; you know you need optical amplification in that case. When I go through that modular circuit, that's electronics; so it's generating power.

I.W.Mackintosh, UK

Could you comment on the potential reliability of the device perhaps measured in the total number of deformations possible on any one element?

Author's Reply

Do you mean reliability in time or in the number of deformities in the initial device?

I.W.Mackintosh, UK

Well, ultimately over its life you're going to expect to deform any one element a certain number of times. Can you comment on what that is likely to project to?

Author's Reply

This is still very much an experimental device, and I hope I didn't give the impression that these are available off the shelf. There are some 128 by 120 arrays that are being experimented on. What is being worked on now is a small part of a larger array to show the density and the addressing can be done for the larger array. People have built these devices in the past on several occasions — IBM, Westinghouse — and at that time the technology wasn't available to place all the addressing circuits in such a small region the size of the mirror. Some changes in the technology and some ideas developed at T.I. have made it possible to produce the device. I think being a silicon based device — well, I don't know how consistent — for this sort of computation I am doing here, if it changes in time it is probably possible to read what.. you could do a test on the device and determine its sensitivity — the brightness relative to the voltage that has to be applied and then make the correction. You know I am using electronics to provide the inputs, so I am essentially using a stored curve for the transducer elements. But I really don't have a good feel. The feeling is that the thing should be pretty good, but I don't have any figures on it.

9. Whitehouse H.T, Speiser J., Aspects of signal processing with emphasis in underwater acoustics. Vol.2, Ed. Tacconi, Reidel, Hingham, Mass. 1977

20. Athale R.A., Lee J.N., Optical processing using outer product concepts. Proc. IEEE, Vol. 72, No. 7, July 1984, p931-941

21. Psaltis D., Two dimensional optical processing using one-dimensional input devices. Proc. IEEE, Vol. 72, No. 7, July 1984, p931-941

22. Lee S. H., Optical analog solutions of partial differential and integral equations, Opt. Eng., Vol. 24, No. 1, jan., 1985

23. 20 Casasent D., Acousto-optic linear algebra processors: Architectures, algorithms and applications. Proc. IEEE, Vol. 72, No. 7, July 1984, p831-849

24. Athale R.A., Collins W.C., Optimal matrix-matrix multiplier based on outer product decomposition. App. Opt. 21, 2089, 1982

25. Athale R.A., Collins W.C., Stilwell P.D., High accuracy matrix multiplication with outer product processor. App. opt. Vol. 22, No. 3, Feb. 1983, p368-370

26. Psaltis D., Casasent D., Carlotto M., An iterative optical processor for adaptive phased array radar processing. Real Time Sig. Processing III. T. Tao, Ed., Proc. SPIE, Vol. 180, April 1979, pp.14-120

27. McAulay A.D., Finite element computation on nearest neighbor connected machines, Symposium on Advances and Trends in Structures and Dynamics, NASA Langley Research Center, Oct. 22, 1984.

28. Caulfield H.J., Floating point optical matrix calculations. Opt. Eng. Vol. 2, No. 6, Nov. 1983

numbers back to decimal is not required until these iterations have been completed. For the 1 billion operation per second machine, with cyclic reduction applied twice, this involves approximately 6000 iterations. Consequently, it takes approximately 48 seconds to solve the equations for a problem of mesh size 500 by 1000. Therefore, the conversion time is negligible.

CONCLUSION

A relatively simple optical computer was described which is capable of performing one billion operations per second with 32 bit data for nearest neighbor computations. It assumes the availability of deformable mirror devices having 1000 by 1000 elements which can be set in 8 mseconds. A simple system with two deformable mirror devices, one CCD, and some simple electronics, operates at 15.6 million operations per second. Higher performance is achieved by duplicating this equipment because the nearest neighbor and residue number concepts permit perfect parallel efficiencies. The matrix operations dominate the computation because overlapping of other operations is possible by algorithm modification.

ACKNOWLEDGEMENTS

I would like to thank Dr. David Thomas and Dr. David Casasent for helpful discussions.

REFERENCES

1. McAulay A.D., Nearest neighbor hybrid deformable mirror optical computer, Proc. SPIE Real Time Signal processing VI, Vol 431, 1983

2. McAulay A.D., Numerical method for adjusting magnetic arrays for medical NMR imaging, IEEE Trans. Vol. Mag-19, No. 5, 1983.

3. McAulay A.D., Prestack inversion with plane layer point source modeling, (Submitted) Geophysics, Jan. 1985.

4. McAulay A.D., Importance of shear in plane layer point source modeling, 54 Annual Intl. Mtg. Soc. Exploration Geophysicsts, Dec., 1984.

5. McAulay A.D., Least square inversion: A novel approach, 53 Annual Intl. Mtg. Soc. Exploration Geophysicsts, 1983.

6. Lee S.H., Ed., Optical information processing fundamentals, Springer Verlag, NY 1981, Ch 4

7. Preston K., Proc. IEEE Intl. Solid State Conf., 100, 1968

8. Preston K., Coherent optical computers, 1972, p 139-148

9. Pape D.R., Hornbeck L.J., Characteristics of the deformable mirror device for optical information processing, Opt. Eng. Vol. 22, No 6, Dec 1983, p 675-681

10. Thomas R.N. et al., The mirror-matrix tube: A novel light valve for projection displays. IEEE Trans., Vol. ED-22, No. 9, Sep. 1975

11. Taylor F.J., Residue arithmetic: a tutorial with examples. Computer, Vol.17, No. 5, may 1984. p50-62

12. Psaltis D., Casasent D., Optical residue arithmetic: A correlation approach. App. Opt. Vol. 18, No2, Jan. 1979, p163-171

13. Jackson J., Casasent D., Optical systolic array processor using residue arithmetic. App. Opt. Vol. 22, No. 18, Sep. 1983.

14. Huang A., The implementation of a residue arithmetic unit via optical and other physical devices. Proc. Intl. Optical Computing Conf., 1975. p14-18

15. Guilfoyle P., Systolic acoustic-optic binary convolver. Opt. Eng. 23, 20, 1984.

16. Alexander T., Computing with light at lightning speeds. Fortune, July 1984, p82-88

17. Kung H.T., Why systolic architectures?., IEEE Computer, Jan., 1982 p37-46.

18. Caufield H.T, Rhodes W.T., Foster M.J.,Horvitz S., Opt. Commn. Vol. 40, No. 2, 1981

APPLICATIONS

Nearest neighbor finite element computations

Equations for solving finite difference and finite element problems with regular, or distorted from regular, grids can be formulated by replacing the derivatives by difference operators. Assume a grid imposed on the field with k a north-south index, increasing in the south direction, and m an east-west one, increasing to the east. The overrelaxation algorithm for updating the k,m th value at the ith iteration in solving an elliptic partial differential equation with a four point difference operator is,

$$u^{(i)}(k,m) = a_1(k,m)\,u^{(i)}(k-1,m) + a_2(k,m)\,u^{(i)}(k+1,m) + a_3(k,m)\,u^{(i)}(k,m-1) + a_4(k,m)\,u^{(i)}(k,m+1) - f(k,m) - \omega u^{(i-1)}(k,m) \qquad (3)$$

The coefficients 'a' have been scaled by the overrelaxation coefficient. f (k,m) represents load that is applied across the grid. Equation (3) is implemented most efficiently by means of a nearest neighbor computer such as the one described. The field u is maintained positive and the whole computation must be performed for positive number coefficients and negative number coefficients separately, as mentioned previously. The field $u^{(i)}(k-1,m)$ is placed on DMD 2, displaced in the south direction by one row as suggested by the k-1 index, and the coefficients $a_1(k,m)$ for the north direction are set on DMD 1. The result of element by element multiplication is accumulated on the CCD. The field for $u^{(i)}(k+1,m)$ is now set on DMD 2, displaced by one row north, and the coefficients $a_2(k,m)$ for the south direction are set on DMD 1. This process is repeated for east, west, load and previous iteration field, while the CCD accumulates the results. The convergence of this approach will be slower than that for a serial machine unless a red-black or odd-even approach is used, (1). In this case, the grid is divided into red and black nodes to look like a checker board. The red points are updated in one iteration and the black on the next. Parallelism is lost unless the CCD is read out on each iteration and the problem mesh is twice that previously used.

Nearest neighbor image processing

The machine is most efficient for operators that are local in nature. Several applications were proposed in reference 1, such as 3 by 3 matrix filtering and averaging over pixel neighbors to compensate for pixel drop out. 3 by 3 filters are commonly used for such processing as edge enhancement. Note that the 3 by 3 filter is applied simultaneously across the whole picture and may vary from one point to another. Image filtering that involves computing values at distances greater than the nearest neighbor can be performed by means of several iterations.

PERFORMANCE

DMD's having 1000 by 1000 elements are anticipated in the next few years. The performance of the proposed system with such a device is considered. The time to set the device in parallel is assumed to be 8 msec. Assuming a moduli set of size 8 in order to produce a 32 bit machine, the DMD's can process an array of size 125 by 1000. 125x1000 numbers are multiplied with the matrix-matrix element multiplier in 8 mseconds, corresponding to a rate of 15.6 million operations per second.

The overall speed need not be much slower than the matrix computation time if sufficient parallel electronics is used in the iteration loop. Perfect parallel efficiency is possible because the modulo operation is completely independent for every value. It is possible to apply a further cycle of cyclic reduction to the red-black algorithm, that is take the red points and impose a further checkerboard on them and repeat this with the black points. This permits the modulo operation on one set of field values to occur in parallel with the calculation of another set, (27). This would be efficient for problem meshes that are twice as large again. One circuit has to perform 1000 modulo operations in 8 mseconds corresponding to a rate of 125 kHz, if there is an electronic circuit for each row of the CCD. Floating point capability is desirable and this has not yet been explored for this processor, (28).

It is possible to increase the machine speed to 1 billion operations per second by simply duplicating the equipment for each residue number modulus (x8) and for each of the eight sums performed in a finite element problem with triangular elements (x8). The multiplications for the nearest neighbor computation for different directions may be performed in parallel with perfect efficiency as suggested in reference 1. The advantage of the residue number system is that each of the code elements may be treated completely independently providing perfect parallel efficiency. $3.72\,N \log N$ iterations of the algorithm are required for an N by N grid, (27). Conversion of the residue

point (FT plane) to filter out low spatial frequencies. The stop is circular for a deformable membrane device and a cross for a cantilever beam deformable mirror. The same lens performs the imaging of the array of dots remaining, the dots having intensities dependent on the mirror deflections.

There is a nonlinear relation between the voltage applied to a drain line of the DMD and the intensity of the spot imaged on the CCD. A look up table is used to convert data values to voltage so as to provide linear intensity of spots with data values for multiplication. As explained in reference (1), only positive values may be multiplied with an incoherent light source. In the operations of interest only one operand will be negative, consequently, the multiplication is separated into two parts, that involving negative numbers and that involving positive numbers. The answers are known to be negative and positive respectively.

A matrix with elements a_{ij} is entered on DMD 1 and a matrix with elements b_{ij} is entered on DMD 2. DMD 1 is then imaged on to DMD 2 and the result imaged on to a CCD. An element of the resulting matrix c_{ij} is the product of the two elements in the same matrix location in the two matrices,

$$c_{ij} = a_{ij} b_{ij} \qquad (1)$$

Element products of other matrices may be added to the result because of the time integrating properties of the CCD. However, the number of matrix-matrix element products that may be added is limited by the dynamic range of the CCD.

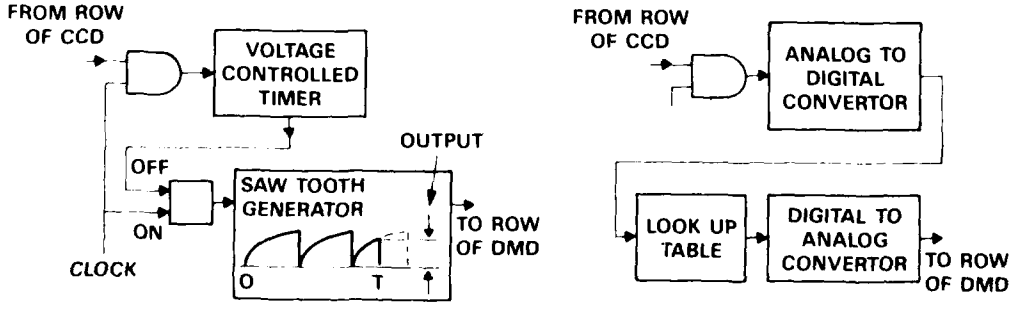
Modulo operation

Multiplying and adding numbers in a residue number system will generate some numbers that exceed the modulos for the number system. The maximum number for a relative prime number in the moduli set is approximately 4 times the product of the relative prime number in the case of a 4 point finite difference scheme. Consequently, it is necessary to reduce these numbers back into the appropriate ranges for the number system. Assume a modulo set m_i, i=1,M and that the result of a computation for the ith member of the set is x_i, then this operation corresponds to computing residuals,

$$R_i = x_i - n_i m_i \qquad (2)$$

where n_i is the number of times that x_i is divisible by m_i.

Figure 6 shows methods of performing the modulo function. Figure 6(a) is an analog system in which the voltage representing intensity from the CCD in the matrix-matrix element multiplier is used to control the length of time for which a saw tooth generator runs. The height of the waveform corresponds to the relative prime number in the modulo set. The generator signal is shaped to compensate for the nonlinearity of the DMD, as described earlier. The height of the generator voltage signal when the timer switches off represents the moduloed value or residue for the specific relative prime number. Figure 6(b) shows a digital modulo system which is likely to be more expensive. A 14 bit analog to digital convertor is required to cope with the largest number which may arise in finite element computations with triangular elements. The look up table provides the corresponding moduloed digital number adjusted to compensate for the DMD nonlinearity. Another method of performing this is shown in reference 13.



(a) Analog (b) Digital

Fig. 6 Electronic modulo and shape circuits

moduli set and keeping the remainder. 7 becomes (1,3,2) because 7 divided by 2 has a remainder of 1, 7 divided by 4 has a remainder of 3 and 7 divided by 5 has a remainder of 2. Similarly 3 becomes (0,3,3). The sum 7 + 3 has the code (1+0, 3+3, 2+3), while the product 7 x 3 has the code (1x0, 3x3, 2x3). The results may exceed the moduli values and each term has to be modulo'ed to keep it the correct range separately. Thus the sum (1,6,5) becomes (1,2,1) and the product (0,9,6) becomes (0,1,1). I refer to this as a modulo operation. Note that no carries were required between elements of the code, unlike fixed radix arithmetic. This enables parallel operation by avoiding sequential operations required for carries. It is necessary to convert back to decimal after completing the computation.

The first reason for using RNS rather than digital multiplication by area, convolution or some other technique to extend dynamic range is that I wish to use the dynamic range already available in the DMD to reduce the number of DMD elements required to represent a number. 5 bits, available on a DMD, permits numbers up to 2**5 = 32 for each of the code numbers in the residue number code. Consequently, numbers may be represented in a residue number system up to 31 x 31 x 13 x 17 x 19 x 23 x 29 x 31 = 6.7 x 10**8 which is greater than 2**32. These 8 relative prime numbers require 8 elements of the DMD as distinct from 32 that would be required for a 32 bit binary convolver. The incorporation of the residue number code into a vector for processing is illustrated in figure 3.

Secondly, the residue number approach permits numbers to be multiplied in a straight forward manner by imaging light from a DMD element, set by the first operand, on to a second DMD element, set by the second operand. The numbers out of the matrix element multiplier are moduloed back to the relative prime numbers for the moduli set selected. No other operations are required with residue numbers because they may be processed independently of each other. In contrast, the systolic acousto-optic binary convolver (15), involves more complexity. It requires coherent light and high frequencies to drive the acousto-optic cell. It requires a sequential shift and add after moduloing the numbers back to binary, following a computation. This is a bottleneck which restricts the capability of the optics.

Outer product multipliers, (20,24,26), require very large dynamic range for the receiving time integration device and this is a major limitation. In addition, the idea of using binary numbers in a 32 bit outer product multiplier is also complicated because it requires the addition of the diagonal elements for each 32 by 32 submatrix, (25). This involves extracting and adding many sets of 32 numbers. This complication is totally avoided by using residue numbers.

Matrix-matrix element multiplier

Figure 5 shows the matrix-matrix element multiplier. A lens is used to image DMD 1 on to DMD 2. It is necessary to use Schlieren type imaging to remove the reflection from inactivated mirrors and from the flat mirror surface between elements. This is accomplished by placing a stop at the lens focal

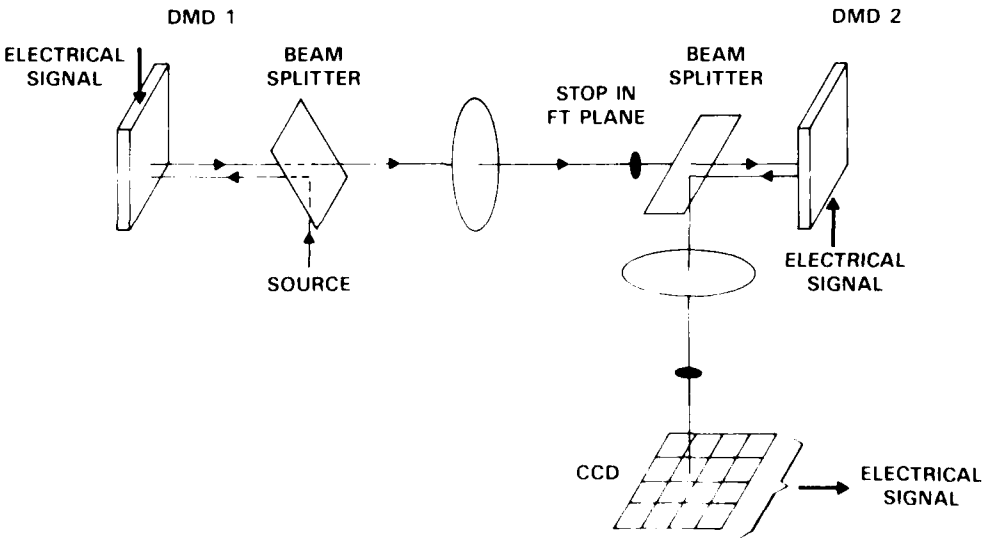


Fig. 5 Matrix-matrix element multiplier for nearest neighbor computation

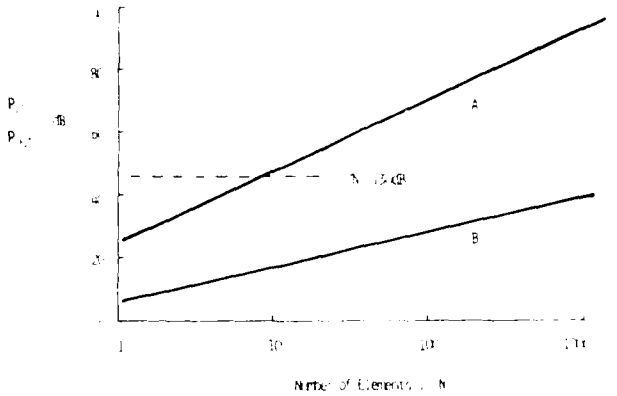


Fig. 10 Ratio of R.F. Power at Manifold Input to R.F. Power at Array Element Input as a Function of Number of Array Elements.
A: Fibre Optic Network
B: Coaxial Cable Network
(typical values)

A fibre-optic link has been used to demonstrate the provision of the r.f. input signal to a 3 GHz phased array module[24]. For this module, an input signal was required at a 10 mW level to drive two mixers, one associated with the transmit section phase locked loop control and the other with the receive section downconversion. The laser output, modulated at 3 GHz, was detected in the module and amplified to the 10 mW level in an FET amplifier.

The single fibre optic link for remote positioning of a radar receiver becomes attractive when the distance involved exceeds some tens of metres and when the dynamic range allowed by the attainable S/N in the required bandwidth is acceptable.

3.2 Phased Array Radar Calibration Systems

An important aspect of phased arrays is their calibration and monitoring. With the severity of future phase and amplitude control requirements, it is unrealistic to expect that active array modules may be totally under open-loop control; some form of regular in-built calibration and monitoring scheme will be required. Such a scheme (Fig. 11)

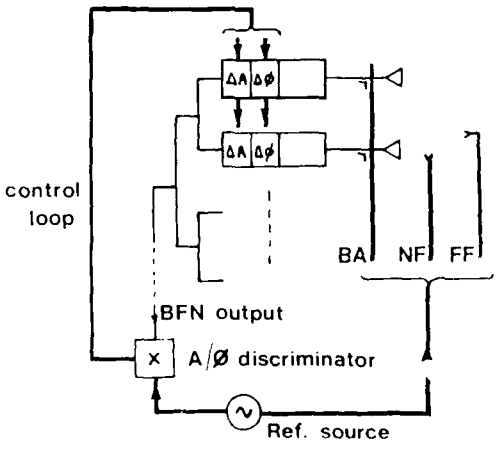


Fig. 11 Phased Array Calibration and Monitoring through Control of Phase and Amplitude Trimmers in Each Element.
FF = Far Field Probe
NF = Near Field Probe
BA = Behind Aperture Probe

requires injection of a calibration signal, either from a far-field source, a near-field source, or a probe in each array element aperture region. Calibration signals passed through the beamformer may be compared with the reference in a phase/amplitude discriminator on an element-by-element basis and control signals derived to adjust phase and amplitude trimmers in each element. The accuracy of the system depends on discriminator accuracy, control loop accuracy and signal injection accuracy, the last-mentioned being the most difficult to achieve; a conventional r.f. manifold to feed a reference signal into each element would suffer from temperature-induced and other path length errors and would severely limit calibration accuracy.

It is therefore proposed that an optical fibre manifold could be used in this application. Use is then made of the high insertion phase stability of this manifold and a single reference signal is distributed to many photodetectors feeding small probes or other suitable r.f. signal injection couplers (Fig. 12). The relatively low level of

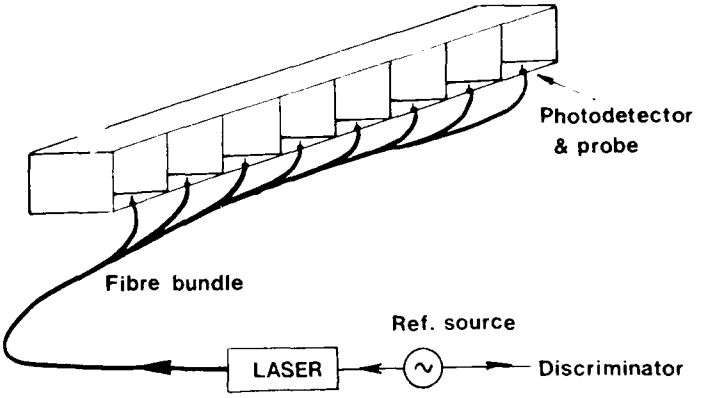


Fig. 12 Fibre Optic R.F. Manifold with Photodetectors for Array Calibration

r.f. signal produced at each detector in such an optical scheme is now not a difficulty since the array receiver, which is part of the system to be calibrated, provides adequate gain.

4. IF AND BASEBAND SIGNAL DISTRIBUTION

4.1 Phased Array Radar I.F. Links

The concern in this application is with the transmission of relatively low level, high dynamic range signals back to the central processor from array elements after downconversion to i.f. at a typical frequency of 70 MHz.

Since signal bandwidths may be 10 MHz or higher, digital techniques, which would be the most common for optical links, imply very high processing rates. For approximately 70 dB dynamic range, 12-bit digitisation would be required at a digitisation rate of at least 20 MHz. This is very much at the limit of current technology, so analogue techniques would be the obvious choice. A natural start would be to consider the use of amplitude, or more strictly, intensity modulation of the optical carrier, followed by baseband, or square-law, detection in a simple photodetector. Using the basic principles outlined in Section 2.3, and presented graphically in Fig. 13, it may be seen

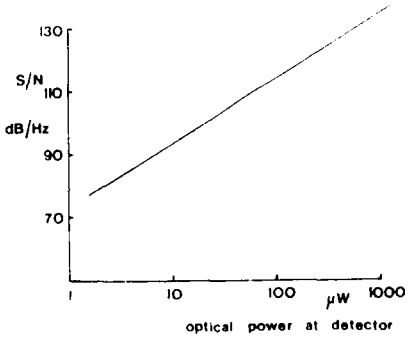


Fig. 13 Typical S/N Performance of PIN Photodetector and Amplifier

that a typical detector will only achieve a value S/N ~ 60 dB in 10 MHz bandwidth at an input optical power of 1 mW - a value close to the point where saturation of the detector occurs. It seems therefore that some 60 dB dynamic range in 10 MH bandwidth is the maximum that can be expected.

Frequency modulation techniques involving applying the signal information to a VCO control input and varying the modulation frequency of a laser, followed by detection and discriminator action, offer slightly higher dynamic range. With wideband f.m. involving a VCO centre frequency of approximately 500 MHz, some 70 dB of dynamic range for signals of 10 MHz bandwidth may be obtained.

The major disadvantages of such schemes which require high dynamic range from small input signals is the need to provide very considerable amplification in each array element in order to provide the required high level of optical signal power at the detector.

4.2 I.F. Beamformers

As indicated earlier, there is growing interest in antennas, particularly for satellite communications applications, that can produce multiple beams. The beamforming network, if used in transmit mode, must take a number of signal inputs designed to be transmitted on particular beams and provide at its output the appropriate coherent signals with amplitude and phase relationships to form those beams (Fig. 14). At

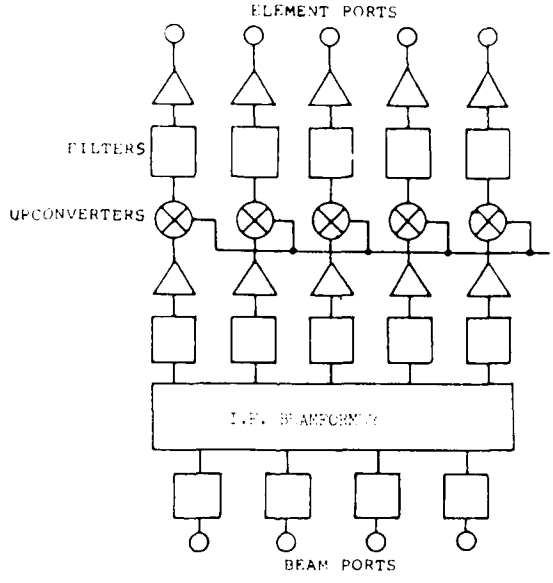


Fig. 14 I.F. Beamformer used with Upconversion to Feed Array Elements

microwave frequency such beamformers are usually configured as complex waveguide structures, or as microstrip Rotman lens structures. For large numbers of array elements or large numbers of beams, these beamformers become very difficult to realise and beamforming at i.f. using resistive matrices is usually preferred . However, such resistive matrix beamformers are limited in their upper frequency of operation and hence bandwidth; a maximum bandwidth of 10 MHz is typical. They also suffer from cross-coupling between ports and somewhat marginal performance in phase and amplitude accuracy (eg. 5-10^0 and 0.5 dB).

A beamformer based on the use of optical fibres as signal delay elements[23] (Fig. 15) to create the required phase relationships at the output ports appears

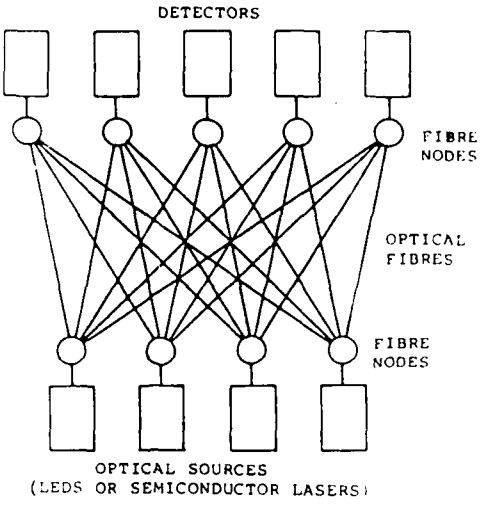


Fig. 15 Optical Fibre Beamformer

attractive because of the high potential phase accuracy and wide bandwidth. The fibre lengths required do not exceed one wavelength at the operating frequency so the beamformer is compact and of low mass. The attainable phase accuracy is limited by the accuracy to which fibres may be cut; this, together with the frequency response limitations of laser diodes and detectors limits the upper frequency of operation.

In communications applications, intermodulation performance is an important consideration. The signal-to-third order intermodulation ratio S/I_3 places constraints on the operating levels of components in the signal path, such as laser and detector, the former usually dominating. The overall performance of the beamformer is summarised in the relation (expressed in dB):

$$B + (S/N) + F + 2L_{EX} + 20\ \log_{10}N_E = 193 + 10\ \log R_E + P_{OI3} - (S/I_3)$$

where (S/N) is the desired signal-to-noise ratio in bandwidth B, P_{OI3} is the third-order intercept point of the laser optical output power, F is the noise figure of the photodetector and amplifier, R_E is the detector load, N_E is the number of elements in the array and L_{EX} represents optical coupling losses (laser-fibre, fibre-detector). For B = 50 MHz and (S/I_3) = 40 dB, a typical system with some 20 array elements achieves (S/N) ~ 30 dB, which matches many communications requirements.

Experiments[25] have verified some of the predictions for such a beamformer and indicate that it represents a very real advance on other beamformers for a number of multiple beam applications. Amplitude control of signals to individual array elements can be provided by neutral density optical filters used as optical attenuators, but a practical problem still remains in finding a good scheme for the splitting of the laser output into many fibre paths.

5. CONCLUSIONS

Optical techniques offer some advantages for signal distribution and control in advanced radar and communication systems.

They are clearly ideal for transporting microwave signals over considerable distances, as in remote positioning of radar receivers, provided high dynamic range is not required and an enclosed transmission path is essential. They are an elegant means of distributing low level r.f. or i.f. signals around an active phased array where these signals are of relatively constant amplitude (as in mixer local oscillator applications). However, there is currently a rather restrictive limit on the size of distribution network possible.

They are obviously suitable for distributing digital control signals to phased array modules and confer considerable immunity to interference.

They are less suitable for high dynamic range signals, such as the received radar returns, either at r.f. or when downconverted to i.f. Future developments in coherent optics or in fast optical A/D technology could, however, influence this conclusion.

Currently, the optimum applications for optical techniques appear to be in i.f. beamformers for multi-beam communication satellite systems and in calibration/monitoring systems for phased arrays.

REFERENCES

1. Ethington, D.A., "The AN/TPQ-36 and AN/TPQ-37 'Firefinder' radars", Proc. Int. Conf. 'Radar-77', London, IEE Conf. Publ. No. 155, pp. 33-35 (1977).
2. Overy, E.D. and Neale, B.T., "Martello - A long range 3D radar", Proc. Int. Conf. 'Military Microwaves 84', London, MEPL (October 1984).
3. Mathias, R.F. and Cunningham, N.A., "AR320 long range defence radar", Proc. Int. Conf. 'Military Microwaves 84', London, MEPL (October 1984).
4. King, 1.E., Albanese, A.A. and Buris, H.A., "The AN/TPS-59 and GE592 - a family of solid state radars", Proc. Int. Conf. 'Radar-77', London, IEE Publ. No. 155, pp. 12-15 (1977).
5. Foster, T.M., "The Westinghouse W-2000, an advanced solid-state long range radar", Proc. Int. Conf. 'Military Microwaves 84', London, MEPL (October 1984).
6. Milne, K., "Principles and concepts of multistatic surveillance radars", Proc. Int. Conf. 'Radar-77', London, IEE Conf. Publ. No. 155, pp. 46-52 (1977).
7. Wardrop, B., "Digital bemforming in radar systems - a review", Proc. Int. Conf. 'Military Microwaves 84', London, MEPL (October 1984).
8. Roederer, A.G., "Multibeam antenna concepts for mobile communications", Proc. Int. Conf. on Satellite Systems for Mobile Communications and Navigation, IEE Publ. No. 222, London, pp. 83-89 (1983).
9. Coirault, R. and Kreidte, W., "Multibeam generation at L-Band: a phased array approach", ESA Journal, 4, pp. 319-336 (1980).
10. Forrest, J.R., "Optoelectronics in practice and promise", GEC Journal of Research, 2, No. 2 (1984).
11. Ackley, D.E., "high power leaky mode multiple stripe laser", Appl. Phys. Lett., 39, pp. 27-29 (1981).
12. Peterman, K. and Arnold, G., "Noise and distortion characteristics of semiconductor lases in optical fibre communication systems", IEEE Trans. MTT-30, pp. 389-401 (1982).
13. Thompson, G.H.B., "Physics of Semiconductor Laser Devices", J.Wiley & Sons, Chichester, 1980.
14. Seeds, A.J. and Forrest, J.R., "High rate amplitude and frequency modulation of semiconductor lasers", IEE Proc. 129, Part I, pp. 275-282 (1982).
15. Gee, C.M. and Thurmond, G.D., "Wideband travelling-wave electro-optic modulator", Proc. SPIE Conf. on Optical Techniques for Microwave Applications, Vol. 477, Washington DC (May 1984).
16. Forrest, J.R., Richards, F.P., Salles, A.A. and Varnish, P., "Optical fibre networks for signal distribution and control in phased array radars", Proc. IEE Int. Conf. 'Radar-82', IEE Publ. No. 216, pp. 408-412 (1982).
17. Sze, S.M., "Physics of Semiconductor Devices", 2nd Ed. J.Wiley & Sons, New York 1981.
18. Wang, S.Y., Bloom, D.M. and Collins, D.M., "GaAs Schottky photodiode with 3 dB bandwidth of 20 GHz", IEEE Electron Devices Meeting, Tech. Digest, San Francisco, pp. 521-524 (1982).
19. Forrest, J.R., Richards, F.P. and Perichou, A., "The Microwave MESFET optical detector", IEEE Electron Devices Meeting, Tech. Digest, San Francisco, pp. 529-532, (1982).
20. "Microcoax" UTI Corp Bulletin 724, Collegeville, PA 19426, USA.
21. Cook, J., "Making low-loss single-mode connectors".
22. Blauvelt, H. and Yen, H., "Transmitter and receiver design for microwave fiber optic links", Proc. SPIE Conf. on Optical Techniques for Microwave Applications, Vol. 477, Washington DC (May 1984).
23. Joseph, T.R., Stephens, W. and Chen, B., "Fiber optic r.f. link", Proc. SPIE Conf. on Optical Techniques for Microwave Applications, Vol. 477, Washington DC (May 1984).
24. Ward, C.J. et al, "A high phase accuracy active phased array module for multifunction radars", IEEE Microwave Symposium, Tech. Digest, Dallas, pp. 179-181 (1982).
25. Sheehan, P.G. and Forrest, J.R., "The use of optical techniques for beamforming in phased arrays", Proc. SPIE Conf. on Optical Techniques for Microwave Applications, Vol. 477, Washington DC (May 1984).

ACKNOWLEDGEMENT

This paper draws on a number of projects supported by UK Ministry of Defence, PE (ASWE), USAF Air Force Office of Scientific Research, and the UK Science and Engineering Research Council.

Particularly useful contributions have been made by F. P. Richards, P. G. Sheehan and C. J. Ward of the Microwave Research Unit, UCL.

DISCUSSION

P.T.Gardiner, UK

You mentioned the phase insensitivity of the fibres due to temperature. Could you comment on the phase sensitivity due to the fibre's sensitivity to strain effects, for example in vibration?

Author's Reply

Yes, the effect of the fibres in regard to strain and vibration, one could think in a similar way of course to temperature effects. Any distortion of the length of the fibres would be important in that regard, so if one varied the length of the fibre just as if it were a temperature effect one would indeed see distortions of phase, and so on. Normally the fibres are very good in terms of their response to strain and where this has been particularly important they have been made up into cables with reinforcing members. So I think the answer to your question is: I don't think there's a problem there. There are problems, of course, with regard to vibration acting as a modulation on the fibre if you are working with base band signals. We are here working with signals which are well above the normal range of vibration frequencies; we are talking about well above a megahertz where the amount of energy that can affect the fibre in terms of distorting loads and so on is insignificant.

S.Ritchie, UK

Concerning your heterodyne technique, would you like to say anything about the linewidth requirements of the lasers?

Author's Reply

Of course it is vitally important in such a system that the intrinsic linewidth of the lasers must be less than the loop locking bandwidth, because otherwise the loop could not lock up. Now it is relatively straightforward to attain loop bandwidths of many tens of megahertz. The current system which we use has a loop bandwidth of round about 50 MHz. The lasers which we use are single mode lasers which have a line width of the order of a MHz. So there is no difficulty in obtaining a locking up of the lasers on that basis. And what happens, of course, in the phase lock loop is that one gets a reduction — a very very strong reduction — due to the loop action of the noise close to carrier. So effectively the laser linewidth is narrowed; the output of the beat frequency from the detector is narrowed from the basic laser linewidths.

ELECTRO-OPTIC TECHNIQUES FOR VLSI INTERCONNECT

John A. Neff
Program Manager, Electronic Sciences Division
Defense Sciences Office
Defense Advanced Research Projects Agency
1400 Wilson Boulevard
Arlington, Virginia 22209
USA

SUMMARY

A major limitation to achieving significant speed increases in VLSI lies in the metallic interconnects. They are costly not only from the charge transport standpoint but also from capacitive loading effects. The Defense Advanced Research Projects Agency, in pursuit of the fifth generation supercomputer, is investigating alternatives to the VLSI metallic interconnects, especially the use of optical techniques to transport the information either inter- or intrachip. As the on-chip performance of VLSI continues to improve via the scale-down of the logic elements, the problems associated with transferring data off and onto the chip become more severe. The use of optical carriers to transfer the information within the computer is very appealing from several viewpoints. Besides the potential for gigabit propagation rates, the conversion from electronics to optics conveniently provides a decoupling of the various circuits from one another. Significant gains will also be realized in reducing cross talk between the metallic routings, and the interconnects need no longer be constrained to the plane of a thin film on the VLSI chip. In addition, optics can offer an increased programming flexibility for restructuring the interconnect network.

NATURE OF THE PROBLEM

The rush to pack more and more computing capability into today's electronic systems will soon be radically slowed unless a major change is made in intra-computer communications. Although problems exist in providing communication links at all levels of computing (from chip-to-chip, through board-to-board, and on to machine-to-machine), the most serious limitations are seen to be at the chip and wafer levels. Many conceived algorithms have not transitioned into architectural implementation because they are so heavily interconnect dependent that the necessary LSI or VLSI chips would consist mostly of interconnects. Let us take a closer look at the origins of this interconnect problem (ref. 1).

There are three problem areas for which optics can significantly impact: increasing interconnect delays, increasing requirements for the number of interconnects, and increasing effects of electromigration. The latter problem is a direct result of scaling down the feature sizes in order to decrease the channel length of active devices, thereby reducing the switching delays. As the cross section of the conductors decrease, the probability of failure in the conductor increases due to increased electron bombardment of the conductor atoms. This can lead to a displacement of these atoms which may result in an open circuit.

The other two problem areas mentioned above result from packing more circuitry onto chips to take up any space that was gained during scaling. In other words, the chip sizes are remaining relatively constant despite continuing attempts to scale down feature sizes. The problem of increasing interconnect delays may best be understood by taking a look at the effects of scaling and packing on the RC time constants of the interconnect conductors. If the linear dimensions are all reduced by a factor alpha, the resistance of the conductors will increase by alpha and the capacitance between the conductors and the ground plane will decrease by alpha. This results in the time constant being independent of any scaling effort. The conclusion to be drawn is that the integrated circuit speeds will eventually be limited by the interconnect speeds because they are not scaling down with the gate speeds.

The above analysis was based only on scaling with no consideration being given to the effects of packing. When packing is considered, the limiting effects of interconnect speed become much worse because the interconnects, on the average, are just as long as they were before scaling. This is based on the assumption that the chip size is remaining constant. If one again looks at the expressions for conductor resistance and capacitance and leaves the lengths unscaled, the resistance is seen to increase by alpha squared whereas the capacitance remains constant. This means that the RC time constant increases by alpha squared resulting in the interconnect delays becoming the major speed-limiting factor for VLSI and VHSIC technologies. The relative magnitude of this problem is illustrated in figure 1 (ref. 2) for the various interconnect metals of polysilicon, tungsten silicide, tungsten, and aluminum.

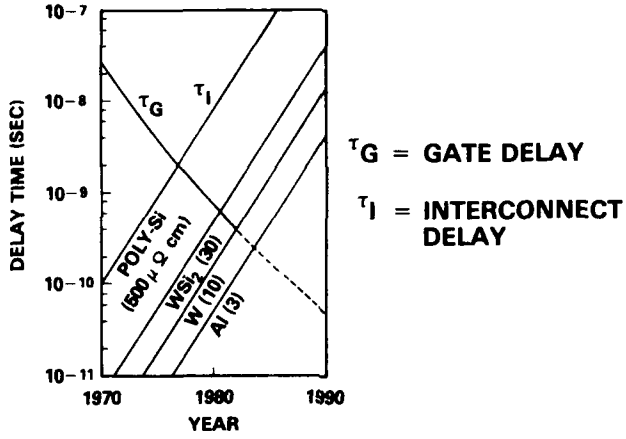


Figure 1. Gate delays versus interconnect delays

The second problem area mentioned above was the increasing requirements for the number of interconnects. According to an empirical relationship known as Rent's Rule, the number of external connections required is approximately equal to the 0.61 power of the number of devices on the chip. This number would be considerably larger for interconnect intensive chips. For VLSI chips with several hundred thousand devices, the interconnect requirements would be considerably in excess of the projected capability (200-300 pins) of the novel solder bump bonding techniques under development. More importantly yet is the limitations that the limited pin counts place on the architectural design of VLSI chips.

ALGORITHMIC IMPLICATIONS

Reducing the cost (power, delay, area, etc.) of interconnections will open the door to new architectures and algorithms that heretofore have been impractical. Architectures have been conceived which would result in more than 75% of the chip area necessarily being expended for the interconnections. These so-called wire-limited architectures are awaiting a solution to the interconnect problem. Let us take a look at the class of algorithms which lead to such interconnect intensive architectures.

The least interconnect intensive algorithms would be those dealing with direct point-to-point dependence of the output on the input. An example would be image contrast reversal where each individual output point is directly related to an individual input point, and each input/output point pair has the same functional relationship. The next higher level would involve neighborhood dependence of output points on the input. Image edge detection algorithms would fall into this class since each output pixel is a function of only those input pixels in the immediate neighborhood of the point of interest.

The above algorithmic classes do not have high interconnect costs. However, the next level of interconnect intensive algorithms deal with more global types of input/output dependence such as is the case with the Fast Fourier Transform (FFT) as shown in figure 2. An important point to note with respect to the FFT interconnect diagram is that all but the first nodal stage could be eliminated if it were possible to reconfigure the interconnects. This would offer a significant reduction in either logic elements or memory transfers for large FFTs. The possibility of realizing such dynamic interconnects will be discussed later in the paper.

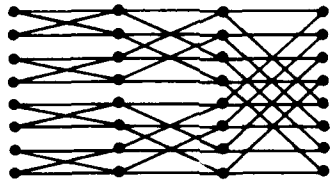


Figure 2. FFT interconnect diagram.

The highest level of interconnect requirements would be presented by space-variant *filtering. Not only are all output points dependent on all input points,* but the functional dependence changes depending on which input/output pair is being considered.

This requires either an extensive interconnect network or many time consuming routing operations.

OPTICAL INTERCONNECTS

Optics offers the potential for realizing very significant improvements in the speed of the interconnects, possibly as high as 10 to 15 gigabits per second. The much higher speed capability of the optical interconnects over the gate speeds provides the opportunity for multiplexing which, in turn, leads to the ability to accommodate a much larger number of interconnects. Although the optical analogy to electromigration could exist with fiber interconnections, there exists little cost motivation to scale down the fiber dimensions to where such considerations would be a factor. Another very real advantage of optics lies in the decoupling of the electronic circuits which the electronic/optical conversion provides. In addition, there would be a freedom from the planar constraints of the electronic microstrips, and the drive power for the interconnects would no longer be dependent on their lengths. But one of the attributes with the greatest potential is the reprogrammability that unguided, or free-space, optical interconnects could provide. Recall the cost savings that such dynamic interconnects would achieve for the FFT.

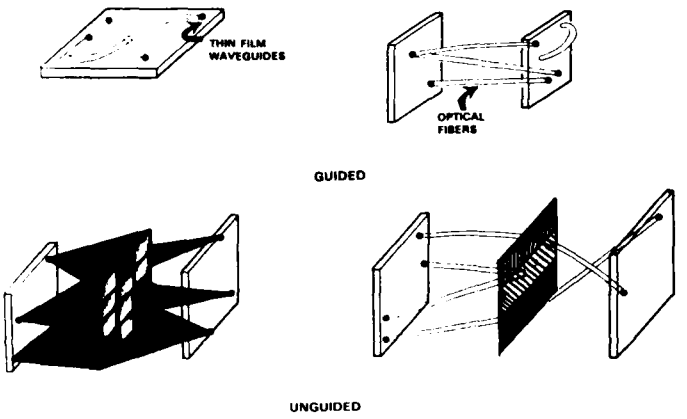


Figure 3. Optical interconnect schemes

Let us take a look at how optics might be used. Figure 3 illustrates four possible concepts - the top two using guided channels while the bottom two are unguided schemes; i.e., they do not involve waveguides. The planar waveguide configuration could be realized by incorporating a waveguide layer in the VLSI chip, but we would still be faced with the planar constraints, and waveguide scattering might generate some cross talk. The optical fiber scheme overcomes these disadvantages, but does not allow us to take advantage of the spatial multiplexing capability of optics. This leads us to look to unguided techniques, which take advantage of the fact that any number of optical beams may independently exist in the same spatial location in a linear medium. The scheme shown on the left employs a mask which determines which sources "see" which detectors. Even if the masking structure were to employ imaging elements which could refract the light to the appropriate detectors, the overall system is not light efficient due to the absorption by the mask. The final concept employs holographic optical elements to diffract the light to the correct location, and affords a great deal of architectural flexibility. An intermixed arrangement of VLSI chips and hologram arrays could someday eliminate the need for the massive bundles of wires currently used to move information inside computers.

Before discussing the reconfigurable interconnects, let us take a look at some existing programs designed to provide optical interconnect technology for the Strategic Computer (fifth generation computer) Initiative of the United States Department of Defense. Figure 4 illustrates an upcoming demonstration designed to study the trade-offs between the optical fiber and the conventional microstrip interconnect technologies. The demonstration will employ 4 emitter-coupled-logic (ECL) word generators each operating at 250 megabits per second. These four circuits will output into a 4:1 electronic multiplexor, two of them via microstrip and two via optical fiber. The output of the multiplexor will then be used as input to a 1 gigabit per second optical fiber channel. For the first demonstration, the silicon based signal generators will be wire-bonded to the GaAs optoelectronic circuitry. Later demonstrations will be moving toward GaAs monolithic circuits and toward more efficient GaAs/Si interfaces.

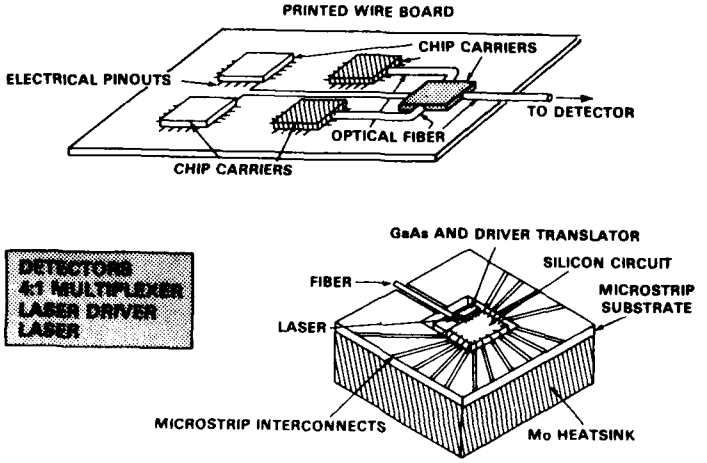


Figure 4. Optical chip-to-chip demonstration

This demonstration will use lasers which emit light horizontally from the edge of the chip, which is the simplest mode of operation for integrated solid state diode lasers. Since this requires that the interconnect signals be brought to the edge of the chips (even for GaAs monolithic integrated circuits), the constraints imposed would lead to very costly interconnects in such situations as inter- or intra-wafer communications. Therefore, a second program will employ distributed feedback lasers (figure 5) monolithically integrated on 3 inch GaAs integrated circuit wafers. The distributed gratings serve a dual function of laser cavity feedback and vertical diffraction of the laser emission. Such surface emitting lasers open up many new architectures, especially when free-space transmission is considered. For example, inter-wafer communications would be possible across a narrow air gap between two wafers mounted upside down with respect to one another. As mentioned earlier, holographic diffraction gratings could be used for inter- and intra-wafer free space links when surface emitting lasers are used.

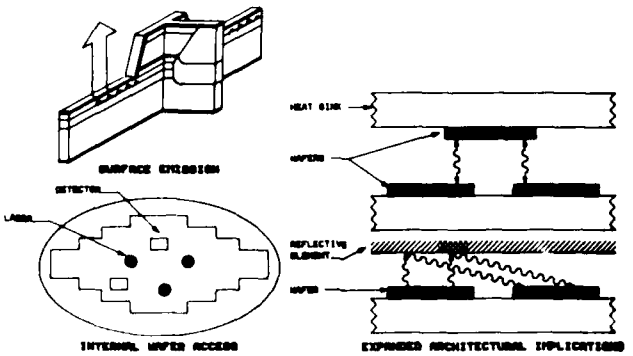


Figure 5. Optical wafer-to-wafer interconnect

Multiplexing is a must if full advantage is to be made of the optical interconnect speeds. High speed electronic multiplexors are costly in terms of chip area, especially for large multiplex ratios. The third optoelectronic interconnect program under Strategic Computing will develop an optical time division multiplexing scheme as shown in figure 6. A small gap is made in each microstrip and the gap is bridged by an optoelectronic material whose resistivity drops by several orders of magnitude upon exposure to coherent radiation. A single laser pulse may be transmitted to each gap (or optoelectronic switch) via differing lengths of optical fiber such that the switches are closed in successive fashion achieving a time multiplexing of the signals on the microstrip. Once again, multiplexing is possible due to the ability to pulse the laser at much higher speeds than the switching speeds on the VLSI chip.

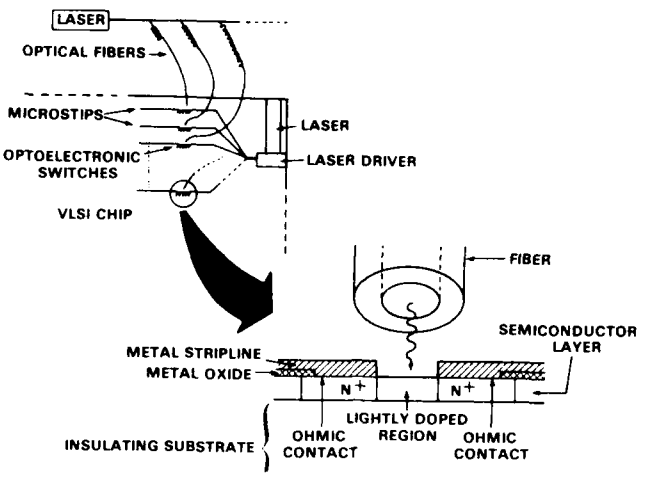


Figure 6. Optoelectronic multiplexing

RECONFIGURABLE INTERCONNECTS FOR VLSI

The interconnect methods considered so far have been fixed in the system design, but optoelectronic interconnects have considerable potential for programmability. This extended capability may be considered for both of the unguided schemes shown in figure 3. In the first scheme, the mask could be replaced with either a two-dimensional spatial light modulator operating in a binary mode or with a two-dimensional array of optical bistable switches. The use of these devices would allow the real-time modification of the mask so that the interconnect configuration could be changed as fast as the logic switching speeds. In the second scheme, the holographic diffraction gratings could be rewritten in real-time using four-wave mixing.

The importance of four-wave mixing as a switching mechanism for optical interconnection and, in general, optical computing, deserves some further discussion. The similarity between degenerate four-wave mixing and conventional holography is illustrated in figure 7 (ref. 3). If one uses the grating point of view, four-wave mixing can be viewed as the simultaneous recording and reading of two sets of gratings. The object wave interferes with each of the two counter-propagating input waves (E_1 and E_2) producing the two sets of gratings. For each set, the non-interfering wave is Bragg diffracted by the respective grating to produce a phase conjugated component wave. These two components then coherently superimpose with one another to produce the phase-conjugate return wave. In this way, four-wave mixing can be used to write the holographic grating structures in a nonlinear optical medium.

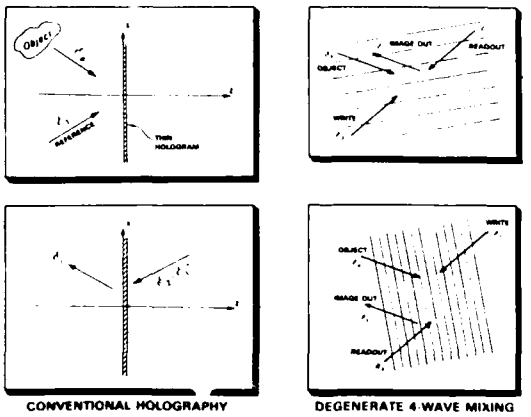


Figure 7. Four-wave mixing for real-time holography

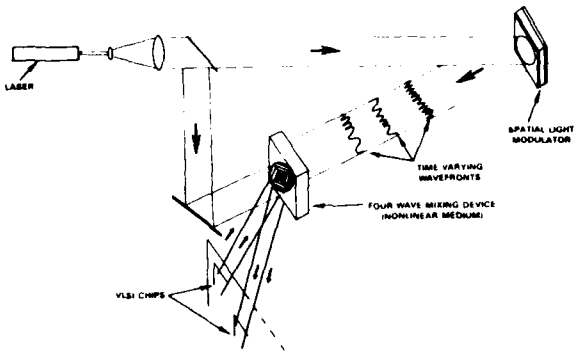


Figure 8. Holographic element arrays using four-wave mixing

For realizing dynamic optical interconnects, large arrays of holographic optical elements would be needed. A viable system configuration for this is shown in figure 8. The laser provides the higher-power counter-propagating waves; however, one of these waves has now been spatially modulated in order to define the various facets of the holographic array. Each facet would be needed to diffract the laser beam of a particular laser diode on a VLSI chip to a particular detector either on the same chip or on another chip. A moment later, the spatial light modulator may impress another modulation on the two-dimensional wavefront, thus establishing a different set of facets in the array. This will result in the light beams following a different path, interconnecting different sets of sources and detectors.

A major impediment to implementing such dynamic interconnect systems lies in the lack of optical nonlinear materials which exhibit their nonlinearity at practical power levels. Both optical bistability and four-wave mixing are dependent on optical nonlinearities, principally the third order nonlinearities. It is much higher magnitudes of these nonlinearities that we must seek if we are to achieve lower power levels. Two emerging areas of nonlinear optical materials research that show exciting promise are superlattice films and organic polymers, both of which have mechanisms for enhancing the third order polarizability. It should also be mentioned that these materials exhibit switching speeds about two orders of magnitude above those of the inorganic films since the predominant switching mechanism is electronic in nature rather than being tied into the lattice vibrational modes.

Nonlinear optical superlattices are constructed by laying down alternating thin films of two semiconductors of differing band gaps, such as alternating films of GaAs and GaAlAs. Electrons produced in the lattice with the higher band energy are rapidly swept into the neighboring layers because of their lower potential. This forced separation of the electron/hole pairs leads to the enhanced polarization that we are seeking. For the organics, the enhanced polarizations are due to the fact that the electrons resulting from the optical excitations are the pi electrons, which are orthogonal to the interatomic bonds, rather than the sigma electrons associated with the bonds. The result is that the pi electrons are more polarizable because they are freer from the interatomic forces which tend to bind the sigma electrons. In our strive for highly polarizable materials, especially those with third order effects, we will also want to investigate the potential of organic superlattices to build on the effects of the superlattices and the organics considered separately.

SUMMARY

This paper has described the interconnect problem and ways in which optics will likely be used to significantly reduce interconnect costs. But not even for the more near term fiber optic systems were the full potential of optical interconnections discussed. The point-to-point communications as described could well give way to elaborate switched or broadcast networks handling packet transmissions. Of course, the flexibility expands even further when free-space links are considered. Realization of the full advantage of optical interconnects will likely require the synergetic efforts of the optical device engineers, the computer architects, and the communication and information scientists.

In addition to the ongoing development efforts into chip-to-chip and wafer-to-wafer optical interconnects, the Defense Advanced Research Projects Agency is interested in pursuing investigations into the above mentioned novel interconnect protocols as well as into many aspects of free-space interconnections. For the latter, the emphasis will be on the switching media, new architectures, better materials, and on improved interfaces between the optoelectronics and silicon based electronics. The reconfiguration potential for optical interconnects, and the associated new architectures and algorithms, may someday revolutionize information processing.

REFERENCES

1. J. W. Goodman, F. J. Leonberger, S. Y. Kung, and R. A. Athale, "Optical Interconnections for VLSI Systems," Proceedings of the IEEE 72, No. 7 (1984).

2. K. C. Saraswat and F. Mohammadi, "Effect of Scaling of Interconnections on the Time Delay of VLSI Circuits," IEEE Transactions on Electron Devices, ED-29 (1982).

3. D. M. Pepper, "Nonlinear Optical Phase Conjugation," Optical Engineering 21, 156 (1982).

DISCUSSION

J.R.Forrest, UK

You mentioned the problem of the heat dissipation in the situation where you are mounting lasers on these chips. You presumably have done some calculations for a typical chip as to how much this would change the heat burden. I wonder if you could give some idea, maybe for a typical case of a VLSI chip taking into account its normal power dissipation, what this would add to the power dissipation.

Author's Reply

I can't myself address that situation. Probably the best person would be Ammon Yariv of the California Institute of Technology who has done a great deal of work on these laser diodes, especially the very high-speed laser diodes, or Dr James Carney of Honeywell who is actively involved in the fabrication of an integrated laser multiplexer structure. At the present time we are only going to try to put a single laser on the chip, but eventually we hope to get several lasers on the chip, and power dissipation, no doubt about it, is going to be one of our chief concerns. But I cannot give you the numbers you want at the present time. I'd contact Professor Yariv on that.

R.Klemm, Ge

I must admit I do have sympathy for what you said at the end toward trying to make it more passive, in other words, having modulators. Recall Dr McCauley's paper: he had a chip with a number of tiny little cells which were deformable. If you could somehow combine that in the geometry you want, you've got a very, very nice passive structure.

Author's Reply

We in no way have made the determination that we're going with one versus the other. In the present system we are going to use lasers but in the more advanced techniques, particularly when we get to the point of wanting to do more and more interconnects to the chip we may be forced, due to the heat problem, to look at the modulators.

J.E.Midwinter, UK

What sort of data rates are actually of interest within the next 5 years between one VLSI chip and another on a single channel?

Author's Reply

I guess I did not use any figures during my talk and I apologize for that. The modulation rate of the lasers that we're looking at the present time are in the 2–3 GHz range. It's been demonstrated that you can go 10 to 15 GHz, but that technology is not advanced to the point that we can consider putting it into manufacturing at the present time. I've actually heard predictions that one could, in the next 5–10 years, conceivably go to 30 GHz on lasers. The laser modulation is the limiting factor here — it certainly isn't the transmission across the fibres since that could go up to 100 GHz or more, but the laser modulation certainly is at the present time. So right now we're only going about 2–3 GHz, but in the silicon circuitry even though VLSIC circuitry working at 20–50 MHz — that gives us a factor of, say 200 that we could multiplex if we really wanted to.

Unidentified Speaker

We want to do as much communication with those chips as we possibly can. There's a lot of algorithms that are not implementable at the present time because they would require significant levels of interconnect, for example interconnecting CPU's with memories in a very parallel fashion.

Numerical Optical Computing at The Ohio State University

S.A. Collins Jr., Professor
S.F. Habiby, Research Associate
A.F. Zwilling, Research Associate

The Ohio State University ElectroScience Laboratory
1320 Kinnear Road, Columbus, Ohio, 43212.

SUMMARY

In this paper we consider designs for various devices intended to perform numerical operations using a Hughes liquid crystal light valve and the residue arithmetic representation. In the first part we present a numerical optical temporal integrator, intended to take as input a string of numbers and keep a running sum. The second device presented is an optical clock designed to generate optical timing pulses using a liquid crystal light valve. A comparison of binary and residue arithmetic in one application is given. The third part contains a discussion of a numerical optical matrix multiplier. A sketch of the design is shown and the information layout on the light valve and input planes is described.

INTRODUCTION

In this paper we present work on numerical Optical computing at The Ohio State University. The work presented in this paper is characterized by the use of a Hughes liquid crystal light valve[1], (LCLV), the residue arithmetic representation, and optical fiber input and output. The paper will focus on the design of numerical optical computing systems with particular application to an optical temporal integrator, that is a device that takes in a string of numbers and keeps a running sum. A residue temporal integrator will be considered along with the optical clock that is used for timing it. Comparison will be made with a comparable unit using binary arithmetic, and the design of a residue-based optical matrix mulitplier will also be presented.

This work falls into the area of numerical optical computing. Optical computing using the optical Fourier transform is well known[2,3]. In the numerical area advances have been made by Athale[4], and Soffer[5]. Residue arithmetic is well known[6,7,8], and optical computation using residue arithmetic has been investigated[9,10,11,12]. In this paper we present an application of optical computing using residue arithmetic in the design of an optical temporal integrator. This contains both an optical addition unit and an optical storage unit. The arithmetic operation of addition is done using addition of optical spatial frequencies[9], and the storage unit operates using feedback[13].

In the balance of the paper we will first review the concepts of residue arithmetic. In the second section we will consider the overall design of the temporal integrator and in the third and fourth sections we will discuss the optical addition unit and the optical storage unit. The fifth section contains a discussion of the experimental apparatus used for the addition and storage units. The sixth section starts a discussion of the optical clock to be used for timing the temporal integrator. In the seventh section the residue-binary design comparison will be made. The eighth section will contain a description of the optical matrix multiplication unit.

1. RESIDUE ARITHMETIC

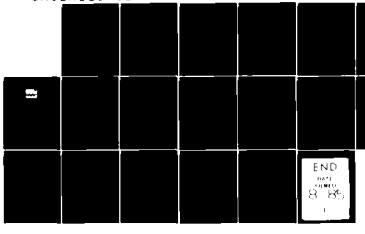
In residue arithmetic, one represents a number by a set of remainders after the number is divided successively by a set of bases. The residue is merely the remainder after dividing the number by the base. Thus, for example, consider the number seven and bases two, three and five. Seven is represented by a residue of one with respect to base two, a residue of one with respect to base three and a residue of two with respect to base five. Thus, the number seven would be represented by the triplet (1,1,2) with respect to bases 2, 3, and 5.

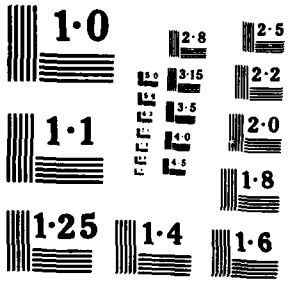
One can perform the operations of addition, subtraction and multiplication quite simply using residue arithmetic. The residues with respect to a given base are combined using the desired operation then the base is cast out if necessary. Consider the residue representations of nine and three with respect to bases 2, 3, and 5; they are (1,0,4) and (1,0,3), respectively. In adding them, start with base two and proceed to bases three and five, adding one to one with respect to base two giving two. One then casts out two giving a remainder of zero. Similarly, with respect to base five, four, the residue of nine, plus three, the residue of three, give seven, which after five is cast out gives a remainer of two. The triplet (0,0,2) corresponds to decimal twelve as expected. Note that all the operations can be performed in parallel.

Similarly, one can subtract and multiply numbers using the residue representation. Subtraction is done by subtracting the corresponding residues, or by adding the complement. The simplest way to multiply is by successive addition. For example, to multiply three by nine, one takes one with respect to base two, no zeros with respect to base three and four threes with respect to base five. The result after casting out two fives is indeed (1,0,2) which represents twenty-seven. Division, if at all possible, is much more complex, resulting from the fact that a larger algebraic field is usually needed to contain quotients than the fields of the dividend or divisor.

AD-A156 906 DIGITAL OPTICAL CIRCUIT TECHNOLOGY(U) ADVISORY GROUP FOR AEROSPACE RESEARCH AND DEVELOPMENT NEUILLY-SUR-SEINE (FRANCE) B L DOVE MAR 85 3/3

UNCLASSIFIED AGARD-CP-382 F/G 20/6 NL




1·0
1·1
1·25
1·4
1·6
1·8
2·0
2·2
2·5
2·8
3·15
3·5
4·0
4·5

It is important to note that for a given base the residues form a cyclic progression. For base five it goes zero, one, two, three, four, zero, one, two, etc. The cyclic behavior is an integral part of the representation. Another important feature is that there is no interaction between the residues of different bases. This saves time in the addition operation for example by avoiding the ripple carry.

2. OPTICAL TEMPORAL INTEGRATOR

We now continue on to the design of a temporal integrator using residue arithmetic. The block diagram of the approach is shown in Figure 1 where we see a clocked input at the left, an adder and two storage units. The output with the running sum is at the right. In operation, the first number that comes in is put in the left-hand storage unit and then transferred to the right-hand one. The next number is added to what is in the right-hand one and then transferred to the left-hand one, etc. The clock pulse isolates the two storage units so that they operate independently yet simultaneously when a new number is being taken in.

3. OPTICAL ADDITION UNIT

The basic addition operation for a single base is shown in Figure 2. It takes in numbers coded by position along a line as shown on the left. There are two inputs, one above the axis and one below the axis. This is consistent with a residue arithmetic representation where all the numbers may be positive. For a residue base, "b", there are "b" possible input positions in both directions. The light valve is shown in the center, and plane P at the right is the output plane.

This residue adder operates by the addition of spatial frequencies. The lens just to the left of the light valve converts the light from the two point sources to collimated beams, i.e. plane waves. These two plane waves interfere to form a fringe pattern on the light valve input face. The average intensity is set so that intensity values lie in a linear region of the light valve characteristic curve. Thus the output intensity, I_1, taken to be proportional to the square of the input intensity, is given by the expression

$$I_1 = C_1 + C_2 \cos(2\pi(x_1+x_2)x/f\lambda). \tag{1}$$

Thus there is a fringe pattern of spatial frequency, κ, given by

$$\kappa = 2\pi(x_1+x_2)/f\lambda \tag{2}$$

where f is the lens focal length, λ is the light wavelength, x_1 and x_2 are the input spot positions along the positive and negative input axis respectively, and x is the position along the light valve input face . Eq. (2) shows the basic addition operation: κ is proportional to the sum of numbers represented by positions, x_1 and x_2.

On the output side of the light valve, one sees a lens with a point source on axis in the focal plane. This provides a collimated beam which is reflected off the light valve. In this case, the periodic pattern incident on the input side of the LCLV acts like a diffraction grating with a cosinusoidal pattern, and the light valve produces three reflected plane waves, one reflected back in the direction of the incoming wave, and one each side of the incoming wave. Upon passing back through the lens these are focussed and appear as three spots in the focal plane, the central spot being coincident with the original source and the other two spots on either side of it. The direction of the other two spots are at angles

$$\theta = \mp \lambda/\ell \tag{3}$$

where ℓ is the fringe spacing indicated in Eq. (1) and is given by

$$\ell = \lambda f/(x_1+x_2). \tag{4}$$

Thus the diffraction angle is

$$\theta = \mp(x_1 + x_2)/f \tag{5}$$

and the separation of the two side spots from the central one in plane P is

$$X = f\theta = \pm(x_1+x_2) \tag{6}$$

thus illustrating that the displacement in plane P is proportional to the sum of the individual displacements. Only one of the two spots is shown in Figure 2.

We now introduce the residue arithmetic by quantizing the allowed values of x along the positive and negative x axes in the input plane:

$$x_1 = nx_0 \tag{7a}$$

$$x_2 = mx_0 . \tag{7b}$$

In Eqs 7, integers n and m are residue representations of an input number with respect to base b, $0 \leq n,m < b-1$, and x_0 is a minimum constant unit displacement. Eqs (1) and (6) thus become

$$I_1 = C_1 + C_2\cos(2\pi(n + m)x_0/f\lambda) \quad (8)$$

$$X = (n + m)x_0 \quad (9)$$

There is more to be done if the output is to represent residue addition and that is to ensure that there is always an output in the range $0 \leq X < (b-1)x_0$. For example, if the two input numbers add up to more than the base then provision must be made for casting out "b" to reduce the number to the desired range. One method is presented in Reference 9. Another will be given here.

The scheme for ensuring an output in the desired residue range is shown in Figure 3. The apparatus arrangement is the same as that in Figure 2 up to plane P. Immediately past plane P a half wave plate is inserted in front of the spots in the range from b to 2b-1. At a focal length from plane P is a lens for collimating the light in the spots and past the lens is a Glan-Thompson prism. Above and to the right of the prism are quarter wave plates and mirrors and below the prism is another lens to refocus the light.

The purpose of the half wave plate next to plane P is to allow spots in the range b to 2b-1 to be returned to the range 0 to b-1 as required by the residue representation. Light coming out in spots in the range 0 to b-1 has unchanged polarization that allows it to pass straight through the Glan-Thompson prism while light coming out throught spots in the range b to 2b-1 has a polarization that is rotated by ninety degrees so that it is reflected by the prism. The field lens next to plane P insures that all the light leaving plane A is directed into the collimating lens next to the prism. The quarter wave plates between the prism and the mirrors effectively act like a half wave plate when light passes through them twice, and rotate the polarization plane so that light initially passing through the prism ends up being reflected by it and conversely, light initially reflected by it ends up being transmitted.

We can now trace the complete light paths. Light from points in the range 0 to b-1 is transmitted through the prism, reflected directly back by the mirror to the right of the prism, reflected downward by the prism and refocussed by the lens. On the other hand, light from spots in the range b to 2b-1 is reflected upward by the prism reflected by the tilted mirror at the top, passes through the prism and spots originally in the range b to 2b-1 now appear in the range 0 to b-1.

This concludes the discussion of the adder. We note that the scheme for reflecting the residue output into the range 0 to b-1 has the advantage of keeping all the light. We also note the preceding discussion considered only one base. Other bases would be in other planes parallel to the plane of the paper and possibly in the same plane but translated sideways.

4. OPTICAL STORAGE UNIT

We now go on to consider the storage units used in the temporal integrator. A single storage unit is shown in Figure 4. There we see the light valve at the center with input at the left and output at the right. There is also an additional path at the bottom feeding information from the output back to the input.

We note that the two lenses on either side of the light valve and the image planes one focal length away are the same as in the addition unit. For the sake of discussion we will denote these focal planes as Q and P in Figure 4. The numbers are position coded as in the addition unit.

Light from spot "n" in the input plane at the left is imaged through the Glan-Thompson prism onto plane Q. There it is added to a second input at the zero spot in plane Q to form the sum, (0+n), giving the number "n" at plane P. Basically plane P could then then be imaged back onto plane Q so that it becomes superimposed onto the original input at plane Q. The Glan-Thompson prism is used because it allows light from the input and from plane P to be combined without the loss of light that would occur in a regular beam-splitter.

The lenslet array shown in the center of the bottom path is added to give the positions a spatially stable equilibrium. Without it there is neutral stability. That is, any position that is fed back will be a stable position. If a small perturbation causes a spot to move, it will remain at the new position. This is shown in Figure 5a where a plot of input position, say in input plane Q, versus output position, say in output plane P, is drawn assuming the lenslet array at the bottom of Figure 4 were absent.

However, with the lenslet array inserted, motion in one direction is corrected by the magnification of -1 of the individual lenslets in the array, thus preserving the spot positions. The lenslet array is shown at the bottom of Figure 4 where the input and output planes are set to give a magnification of -1. The input versus output positions are then redrawn in Figure 5b where we see a separate portion for each lens and a negative slope for each portion. The equilibrium spot positions occur where the individual segments cross the slant line, where input position coincides with output position, as shown. The intersections where the slopes are opposite give stable equilibrium. It is reasonable that a stable equilibrium should occur because the axes of the individual lenslets determine unique spot positions.

The lenslet array is formed from a bleached holographic sine-wave grating. The grating is formed from the photograph of two interfering plane waves. The refractive index of the photograph, when it is bleached is given by $n = n_0 + n_1\cos(\kappa x)$ where x is the position coordinate on the film perpendicular to the fringes and $\kappa = 2\pi/\ell_0$ where ℓ_0 is the fringe spacing. The focussing action then occurs at around each lenslet axis given by maxima in refractive index, with the focal length being given by the expression $f = -(2\pi/\lambda)(d^2n/dx^2)$ where λ is the optical wavelength.

In order to ensure interference at the light valve input, precautions must be taken. The polarizations of the individual beams must be carefully set. The same source must be used for the light at the input planes and at the two position planes. One way of achieving this would be to use optical fibers all coming from the same source all with the same path lengths.

We now combine the adder and the storage units into the completed temporal integrator shown in Figure 6. There we see the adder across the top and the two storage units across the bottom. The input is at the upper left and the output is at the lower left. The information travels in a clockwise manner around the loop. The prism used to cast out the base is located at the upper right hand side. The output of the addition unit is applied directly to the input of the first storage unit, the output of which is applied directly to the second storage unit. Clocked sources are used at particular inputs to disconnect the two storage units during the input of a new number.

In the temporal integrator, information is partitioned in rows. That is, the individual gratings on the light valve are only one image element high. This is illustrated in Figure 7 where we see the light valve and input plane and two lenses, the spherical lens described previously and, in addition, a cylindrical lens. The cylindrical lens serves to image in the vertical direction, leaving the light to be collimated as before in the horizontal direction. This has the obvious advantage of independent rows.

5. EXPERIMENTAL APPARATUS

For clarity and simplicity, the apparatus is illustrated in two parts, the adder and the storage unit. The addition unit is shown in Figure 8 where we see the light valve in the center and the laser source at the top. A twenty-five milliwatt single line argon-ion laser was used. The light is first split into two parts, one for viewing the output and one for the input. The input beam is then split into two more parts. These go through two fibers, one from a set of fibers that indicate the value of one input number and the other from a set of fibers that indicate the value of the second input number. There are several fibers positioned to allow several different values for the each input number. These are accessed individually by directing the input beam into one or another fiber. The fiber ends are then positioned next to one another on a single fiber-holding block as shown. Only three fibers are shown in Figure 8; the actual addition unit has twenty-two fiber inputs for base 11 residue arithmetic addition. The light from the fiber-holding block goes through a spherical-cylindrical lens combination and is imaged to a slit-shaped interference pattern on the light valve input side.

On the output side of the light valve the primary beam comes in from the right and is imaged to a point on a mirror in an image plane to the right of the light valve. It then passes through a spherical-cylindrical lens combination onto a slit-shaped pattern on the light valve. The reflected light then passes back through the combination lens and is reimaged back to the position plane . Half of the position plane is covered by the mirror and the other half is partially covered by a half-wave plate. The reflected light then diverges and is collimated to pass through the Glan-Thompson prism at the lower left. It is then reflected off the two mirrors outside the Glan-Thompson prism and reimaged, giving the final output at the right of the prism.

The storage apparatus is shown in Figure 9. The light valve is just to the right of center with its reflective output face showing. The circular configuration of the light path is shown. The read beam comes in from the lower left through a pinhole and a Glan-Thompson prism and a lens to the light valve. The desired portion of the light reflected from the light valve is again reflected by the polarizing prism and two successive corner mirrors, one of which will be semi-transparent. At the upper right it is reflected again by a polarizing prism back onto the input surface or the light valve. The numerical information is applied to the loop through the polarizing prism at the upper right and the output will be taken through the semitransparent mirror. The lenses in front of and in back of the light valve are spherical-cylindrical lens combinations (designated elliptical in Figure 9) so they provide horizontal slit-shaped patterns on the light valve. The other two are spherical Fourier transform lenses.

In operation, the information comes in from the upper right in the form of light from two point sources. These are collimated in the horizontal plane and imaged in the vertical plane to the slit-shaped input pattern. The light from the read source comes in through a small aperture and the polarizing prism at the lower left and is imaged by the "elliptical" lens to the corresponding slit shaped pattern on the output face of the light valve. The appropriate polarization component of the reflected light is imaged to two points on the holographic grating lenslet array to maintain the positions of the spots. One of the two spatial sidebands is blocked at the grating and the light from the other two spots passes through the two Fourier transform lenses to be reimaged at the center of the side opposite the grating. It then passes through the polarizing prism and elliptical lens to be reimaged into the slit-shaped grating on the input side of the light valve.

This concludes the section on the residue arithmetic temporal integrator. The optical clock is considered next.

6. OPTICAL CLOCK

We now go on to consider the optical clock design. As indicated this will be similar in design to a simple electronic multivibrator. We will consider first the electronic version, then schematically indicate an optical implementation. Finally an experimental configuration will be indicated.

The electronic circuit equivalent to the optical clock is shown in Figure 10. There we see two transistors in a bistable configuration. In addition there are delays represented by the RC circuits feeding back from emitter to base. The key items are the bistable unit and the delay elements.

The equivalent optical circuit is shown in Figure 11a. This is divided into two parts shown in Figures 11b and 11c. The basic bistable flip-flop is shown in Figure 11b. There we see two light valves configured so that the output of one serves as the input to the other and vice versa. In actuality the symbol denoting the light valve represents only one image element of a light valve. The prisms marked GT are Glan-Thompson prisms assumed to pass light polarized in the plane of the figure and reflect light polarized perpendicular to it. The light valve AC excitation is set so that with the proper input polarization and the Glan-Thompson prisms acting as analyzers on the output side, a dark input gives a bright output and conversely a bright input gives a dark output. Thus a dark input to the top light valve gives bright input to the bottom light valve which in turn gives the dark input to the top one, providing a consistent situation. If we interchange light and dark then there is another consistent situation, giving the bistable unit.

The delay units are shown in Figure 11c which is a copy of the right hand side of the complete circuit in Figure 11a. There we see two identical loops both having identical operation. In the top one light comes in from the upper left through the Glan-Thompson prism onto the light valve numbered 3. A bright input there causes the light on the output side to go dark after one response time of the light valve. This output is then applied through an analyzer to the input of the light valve numbered 5. After a second response time the output of that light valve past the Glan-Thompson prism then switches to bright.

The essential features of the delay loop is that if the light valve is set up so that the Glan-Thompson prism is reflecting dark from the top light valve in the flip-flop, then after a delay of two light valve response times bright is transmitted through the Glan-Thompson prism. Both of the delay loops in Figure 11c operate identically providing the required delays for both portions of the flip-flop.

The next step is to design an optical configuration with only one light valve incorporating the six image elements shown in Figure 11a. This design is represented in Figure 12. There we see the single light valve in the center with two optical loops, a square one on the bottom and a triangular one on the top.

The square bottom loop is the flip-flop loop. With the single lens, focal length f, positioned a distance 2f from the light valve, a point on one side of the axis on the output side, represented by the dot, is imaged onto the input side on the other side of the axis, as represented by the second dot. The output from that point is then imaged on the first side of the axis, onto the input corresponding to the original point.

The delays are provided by the triangular loop. Both spots in the pair are imaged simultaneously around that loop, one spot corresponding to the top delay loop in Figure 11c and the other point to the bottom delay loop. Light from the two spots on the output side of the light valve, labled 1 and 1' in Figure 11a is thus imaged by the lens in the right hand side of the triangle onto the mirror at the top of the loop in Figure 12. The lens on the left hand side of the triangle then images the spots back onto the input to the light valve. The corner mirrors in the triangular loop are slightly tipped so that the spots are reimaged on the light valve input to points below the plane of the figure, points 3 and 3' in Figure 11a. This also means that light coming off the output side at 3 and 3' will be imaged around the triangular loop below the points at 3 and 3' providing a third pair of spots on the light valve input, those corresponding to 5 and 5' in Figure 11a.

The mirrors at the top and to the left of the triangular loop in Figure 12 are segmented to return light in the delay loop to the original flip-flop. The two segments are divided by a horizontal line. The segment of the mirror above the line will be tipped in a direction different from the segment below the line. The pairs of spots imaged onto the mirror at the top of the triangle from the first and second times around that loop, points 3 and 3', and 5 and 5', strike the mirror below the dividing line. On the third pass the pair of spots striking the mirror at the top of the triangular loop hit above the dividing line. The top segment of that mirror then directs the light reflected off it to the bottom segment of the mirror at the left of the loop. That mirror segment is then oriented to direct the light so as to superimpose it on the inputs to points marked 1 and 1', in Figure 11a.

The preceeding is shown nicely in Figure 13, a three-dimensional portrayal of the arrangement of Figure 12 . We see the light valve in the center of the figure. The spots 1, 1', 3, 3', 5, and 5' are shown on the input side of the light valve. To the bottom is the square loop with its one lens. It images spots 1 and 1' each back onto the other. At the top is the triangular loop with its two lenses and segmented mirrors. Spots 1 and 1' on the light valve output are imaged onto spots 2 and 2' on the lower part of the bottom half of the segmented mirror at the top of the figure, which are in turn imaged onto spots 3 and 3' below points 1 and 1' on the light valve input. Similarly points 3 and 3' get imaged to spots 5 and 5'. The light from spots 5 and 5' on the light valve output then goes to spots 6 and 6' on the top segment of the top mirror, and from there to the bottom segment of the left-hand mirror and to spots 1 and 1' on the light valve input.

It is noted that although two passes around the delay loop are shown, many more passes around the loop are possible. One merely raises the dividing line between the two segments on the top mirror.

The output can be taken from any one mirror by making it partially transmitting. Alternatively, the beam transmitted by the right-hand Glan-Thompson prism in Figure 11a would be an ideal port since the transmitted beam is unused anyway.

An actual experimental arrangement is shown in Figure 14, drawn from a photograph of the apparatus. There we see the input to the light valve on the right and other components as indicated in Figure 13. The spots on the light valve are roughly three millimeters in diameter.

As a preliminary indication of the operation of the pulser the oscilloscope trace in Figure 15 is shown. There we see a periodic waveform taken at the output port of the pulser. The pulse rate is rather slow. However, the pulse rate is connected to the response time of the light valve. For faster light valves the pulse rate will be faster. Often in computing operations the clock rate is not that of the fastest operation, so it is not necessarily a problem that the pulse period is several light valve response times in duration.

7. DESIGN CONSIDERATIONS

We now want to present some general design considerations. Specifically we want to compare the design for the temporal integrator using residue arithmetic with a design for a comparable unit using binary arithmetic. We will be interested first in the number of parallel units that can in principle be configured on one typical light valve, and second in the response times of the two configurations. To start with we will first estimate the number of image elements required for a given dynamic range, the number of residue temporal integrators on one light valve, and the response time. Then we will present a temporal integrator design using binary arithmetic and estimate the same quantities.

The number of light valve image elements required can be inferred from Figure 6. There we see the complete temporal integrator configuration for one residue base. We note that for the top light valve there are 2n image elements required in the input plane for that base. A simple calculation will show that to have the output plane spots the desired size, there are also 2n image elements required on the light valve input plane. Simarly for the same base there are n light valve spots required for each storage unit, making 4n LCLV image elements required per base. For a number of bases, call them n_1, n_2, etc. there are $4\Sigma n_i$ image elements required per temporal integrator. For a light valve with 600x600 image elements, a total of N_r parallel temporal integrators is possible with N_r given by

$$N_r = 600^2/4\Sigma n_i . \tag{10}$$

We note that the dynamic range is $R=\Pi n_i$ and that it would require three light valve response times to add one more number to the sum in memory.

We now consider the corresponding design using optical binary logic. The optical logic design is shown in Figure 16. There we see an array of light valves and optical rays. The design is broken up into three parts, sum and carry operations representing an adder and a master-slave accumulator representing the storage operation. We will not dwell here on the details of the operation of the unit. They have been given elsewhere[14]. Suffice it to say that the sum and carry operations are composed of optical AND and exclusive OR operations, and the master-slave flip-flop is composed of simpler flip-flop operations. The figure shown applies to one binary bit. One can imagine that the diagram would be iterated above the plane of the figure as many times as there are binary bits.

The main point is that to determine the required light valve space we note that there are thirteen rays intersecting the input side of light valves, implying that thirteen light valve resolution elements are required per binary bit. Thus for a dynamic range of $R=2^b$, i.e. "b" bit numbers there will be 13b image elements required per temporal integrator, or N_b total temporal integrators possible on the 600x600 image element light valve where N_b is given by

$$N_b = 600^2/13b \tag{11}$$

We also note that there would be 2b light valve response times for each number added to the sum, much larger than the residue arithmetic case.

To compare the two numbers of temporal integrators in terms of range R, we form the ratio N_b/N_r approximately given by

$$N_b/N_r = 4q(R)^{1/q}/13\log_2 R . \tag{12}$$

Eq. (12) applies best when the residue bases, taking q of them are closely spaced.

As one means of comparison we consider the various residue base combinations. Noting that the residue dynamic range is merely the product of the residue bases, we see that it is possible to achieve the same range with a few large bases or more smaller bases. The two different extremes are compared in Figure 17 where we see the dynamic range plotted as a function of the ratio N_r/N_b. We see that for the same number of temporal integrators, the range is larger if five small bases are used rather than three or four larger ones.

Another graphical comparison is shown in Figure 18 where we see plotted two curves as a function of dynamic range showing respectively the number of residue temporal integrators and the number of binary temporal integrators. We note that for small dynamic ranges the number of residue integrators is slightly larger, with the number becoming equal for larger dynamic ranges. Thus we note that roughly speaking the two schemes are comparable. The telling item is the response time. For sixteen bit numbers the residue temporal integrator is a factor of ten faster, mainly due to the ripple carry. These numbers can not be offered as conclusive. However they do provide an indication of how the trends might be expected to go.

8. MATRIX VECTOR MULTIPLIER

We now consider a design for a somewhat different apparatus, namely an optical matrix multiplier. The design uses as in the previous devices a Hughes liquid crystal light valve and the residue arithmetic representation. It further uses position-coded information which is intended to come in through optical fibers. It has the further restrictions that it operates only on matrices that have at most on element per row. The position coding is chosen so that the input format is the same as the output format so such units can be cascaded. The unit is intended to take in a set of numbers representing the values of the components of a vector, call it $\underline{b}$. Each of these components is then to be multiplied by the corresponding elements of a matrix, call it matrix $\underline{a}$, generating the products $a_{ij}b_j$.

In the description of the multiplier, we first present the table-lookup multiplication procedure, then the operating principle. This will be followed by the input and output coding arrangement and a view of the device for a single residue representation will be presented. Finally the approach for extending the scheme to a residue representation with an arbitrary number of moduli will be indicated.

The actual numerical multiplication is performed using a table-lookup mapping technique[9,15]. This is illustrated at the left in Figure 19 where we see two mappings. The inputs for each mapping are on the left and outputs are on the right with the interconnections shown as lines . Mappings for multiplication and addition are one-to-one and thus are unique. Shown in the figure are mappings for multiplication by one and by four for the residue modulus nine. For example seven multiplied by four would be twenty-eight in the decimal number system. In the residue number system, modulo nine, the result is one. Thus in the table, the input of seven leads to the output of one.

The maps representing the tables are also shown in Figure 19. The input is represented along the vertical axis and the output is represented on the horizontal axis. For modulus nine the table has size nine by nine representing integers zero through eight. The spot in each row indicates the connection between input and output. Thus the spot in the row corresponding to input seven is in the column corresponding to an output value of one. The fact that there is one and only one spot in each row is the indication that the mapping is one-to-one.

The basic operating principle for the multiplier is indicated in Figure 20. There we see a liquid crystal light valve on the right with input beams on the left and a special mirror to its right. The mirror is defined for the purpose of illustration and will be replaced with more easily obtainable apparatus later. It has the property that light of one polarization, say the plane of the page, is transmitted through it and light of the other polarization, perpendicular to the plane of the page, is reflected back off it. The light valve has the property that with zero input intensity it leaves the plane of polarization unchanged. With maximum input intensity it acts like a half-wave plate, rotating the plane of polarization by ninety degrees.

In operation, a read beam of light polarized in the plane of the figure comes in from the right, is transmitted by the mirror and is imaged onto the light valve at the position where the top input beam is incident on the opposite side. The light valve thus changes the polarization plane so that the light reflected off the light valve then reflects off the mirror back to a position on the light valve where there is no input spot. With no inputs, the polarization remains unchanged so the beam continues to bounce back and forth between the light valve and the mirror. After the desired number of reflections another spot of light is applied to the input side of the light valve. This again rotates the polarization plane by ninety degrees and the light is transmitted through the mirror. The net result is that the spot is translated a number of units sideways, depending on the position of the second beam on the input side. This procedure corresponds to moving along a line in the maps on the left side of Figure 19.

A simple matrix multiplication scheme using this approach is shown in Figure 21. There we again see the light valve and the polarizing mirror. Light can come in through any one of a vertical column of input positions shown to the right of the mirror; the vertical position indicates the numerical value. The light coming in has a polarization such that it passes through the mirror. Upon the first light valve reflection, the polarization is rotated so that it reflects off the mirror bouncing back and forth between the light valve and the mirror. The number of bounces in a given row is determined by the second light valve input spot, the position of which is determined by the mapping of the type shown in Figure 19. After the light is transmitted through the polarizing mirror, it is directed towards the horizontal axis by the cylindrical lens. The output value is then indicated by the position along the horizontal output axis.

The light for the input from the left of the light valve is intended to come from a microprocessor-controlled CRT. It is assumed that the information on the matrix elements is stored in the microprocessor and needs to be changed relatively infrequently. The light coming to the input plane at the right could well come in through multimode fibers. The multiplication proceeds with the speed of light, with an a new residue product as often as a new input appears in the column of possible input vector positions.

A more practical configuration is shown in Figure 22. There we see the light valve to the left of the center of the figure with an input CRT to its left and optical circuitry to its right. The numerical information input is shown at the upper right and output at the lower right. The light comes in through the Glan Thompson polarizing prism with lenses on each side of it. The prism is one focal length from the light valve and one focal length from the input plane so that the light is collimated in the polarizing prism. In actuality the glass in the prism should be taken into account so that the prism with the two lenses should be treated as a single thick lens with the light valve and input planes in its focal planes. The prisms are intended to pass light polarized in the horizontal plane and reflect light polarized in the vertical plane.

In operation an input spot of light is imaged onto the light valve by the polarizing prism lens at the top of the figure. Only the polarization component passing through the prism is retained. The polarization is then rotated by the light valve so that it is reflected by the polarizing prism in the second polarizing prism-lens combination. It is then reflected off a diagonal mirror and reimaged onto the mirror at the far right. There is a field lens in front of that mirror for light management purposes. The light then goes to a second diagonal mirror, reflects off the top prism-lens combination, and is imaged again onto the light valve. The second diagonal mirror is tipped so that the second spot on the light valve is slightly displaced horizontally thus providing the second spot in the row of spots corresponding to that in Figure 21. When the spot has been displaced the requisite number of positions, an input to the light valve again causes the polarization to be rotated so that it passes straight through the bottom polarizing prism into the cylindrical lens to the output plane as in Figure 21.

A few items might be mentioned about the scheme just described. It is designed for finite size mappings representing full connection or no connection. It works well with any mapping such as those required by the residue arithmetic transformations. It is less versatile than previous matrix-multiplication schemes[16] but has much lower light loss. In that sense it would operate well as the distribution mapping in the computer switching system of Clymer[17].

The preceding has been the multiplication unit for one residue base. In practice a number is represented by three or more residue bases. We now extend the situation to indicate the coding for more than one residue base. One might ask how this fits into the complete matrix-multiplication scheme. The approach is indicated in Figure 23. There we see the complete matrix multiplication unit symbolically presented. The input vector is shown on the left with the vector elements divided into three sets of spots, each representing a residue modulus. The implication is that there would be three units like those just described, one physically above the other. The box labeled 'mapping' contains the mappings and the a_{ij} generator, indicating, as mentioned, that the mapping is generated in the microprocessor controlling the CRT input to the multiplication units. The output of the multiplication units is a set of products, the $a_{ij}b_j$, as indicated. These form the input to an addition unit similar to the one discussed in section 3. As would be expected each $a_{ij}b_j$ product is still represented as the set of three residues. The spots representing the $a_{ij}b_j$ products are shown vertically in Figure 23, symbolically representing the output of a multiplier, although those shown in the output plane of Figure 22 are horizontally arrayed.

REFERENCES

1. W. P. Bleha, et. al., Opt. Eng. 17(4), 584 (1978).
2. J. W. Goodman, Introduction to Fourier Optics, McGraw-Hill, New York (1968).
3. K. Preston, Jr. Coherent Optical Computers, McGraw-Hill, New York (1972).
4. R. A. Athale and S. H. Lee, Appl. Opt. 20, 1424 (1981).
5. B. H. Soffer, D. Boswell, A. M. Lackner, P. Chavel, A. A. Sawchuk, T. C. Strand, and A. R. Tanguay, Jr., in Proc. SPIE 232, 128 (1980).
6. N. S. Szabo, and R. I. Tanaka, Residue Arithmetic and its Applications to Computer Technology, McGraw-Hill, New York, (1967).
7. H. L. Gardner, The Residue Number System, IEEE Trans. Elec. Comp. EC-8, 40 (1959).
8. A. Huang, Digest of Papers, International Optical Computing Conference, IEEE Catalog No. 75CH0941-5C, 14 (1975).
9. D. Psaltis, and D. Casasent, Appl. Opt. 18, 163 (1979).
10. A. Tai, I. Cindrich, J. R. Fineup, and C. C. Aleksoff, Appl. Opt. 18, 2812 (1979).
11. F. A. Horrigan, and W.W. Stoner, Optical Information Systems Processing Conference, Proc. SPIE, 185, 19 (1980).
12. C. C. Guest, and T.K. Gaylord, Proc. SPIE 185, 42 (1979).
13. S. A. Collins, Jr., and K.C. Wasmundt, Opt. Eng. 19, 478 (1980).
14. M.T. Fatehi, K.C. Wasmundt, and S. A. Collins, Jr., to be published, Appl. Opt., July 1984.
15. C. C. Guest, and T.K. Gaylord, Appl. Opt. 19, 1201 (1980).
16. J. W. Goodman, et.al., Opt. Lett. 2, 1 (1978).
17. B. Clymer, Thesis (M.Sc.), Ohio State University, 1982.

ACKNOWLEGMENT

This work was supported in part by Grant No. NSG3302 between National Aeronautics and Space Administration and the Ohio State University Research Foundation.

FIGURES

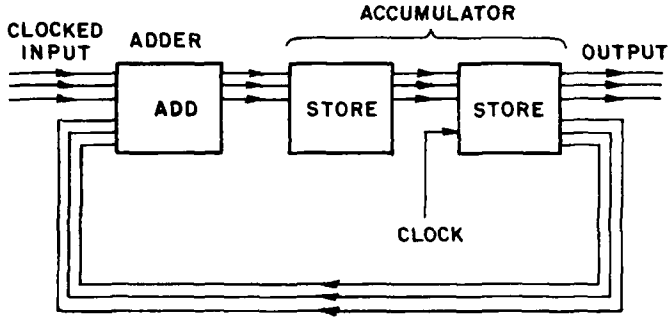


Figure 1. Block diagram of residue temporal integrator.

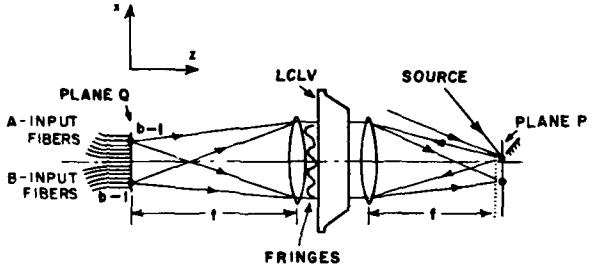


Figure 2. Optical Residue adder

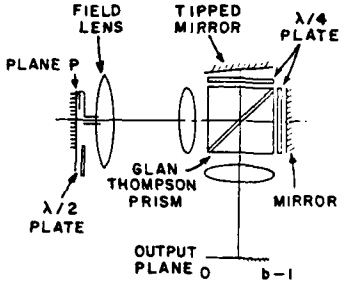


Figure 3. Addition unit base casting mechanism using a Glan Thompson polarizing prism.

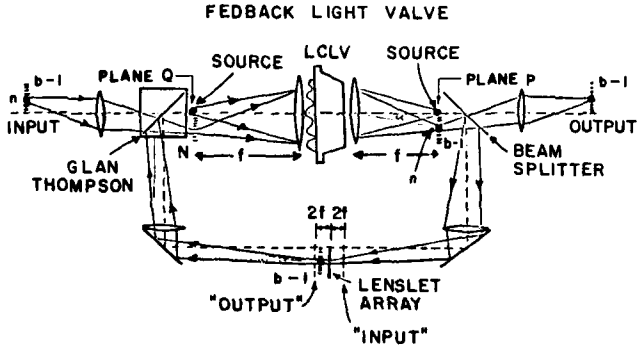


Figure 4. Optical residue storage unit.

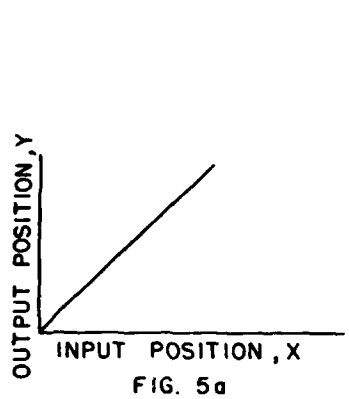


FIG. 5a

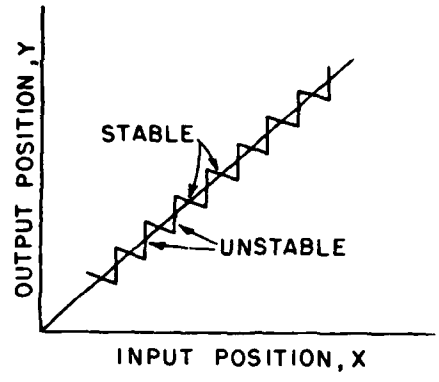


FIG. 5b

Figure 5. Graph showing relation of input to output spots
a. with no lenslets
b. with lenslets.

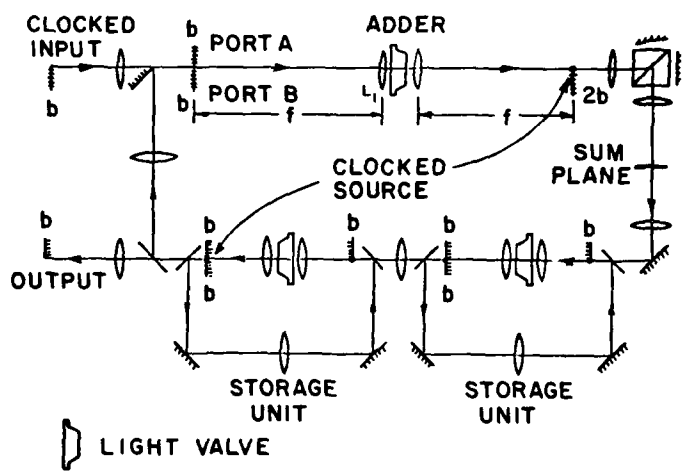


Figure 6. Optical Residue Temporal Integrator.

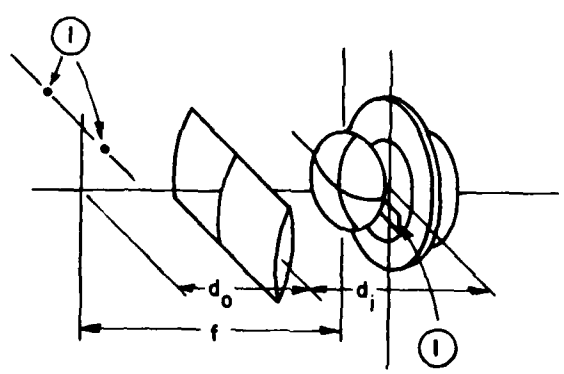


Figure 7. Sketch showing cylindrical lens for vertical imaging.

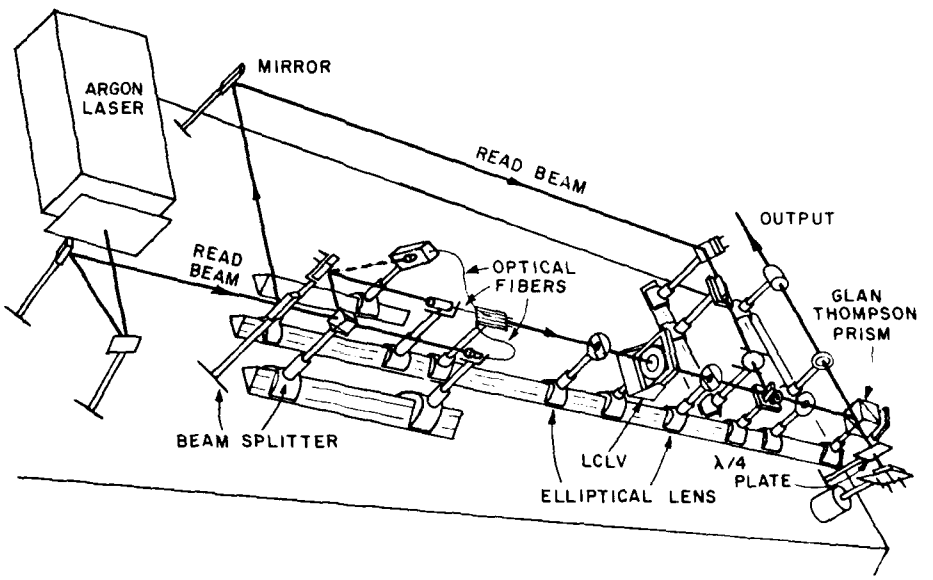


Figure 8. Sketch showing addition apparatus.

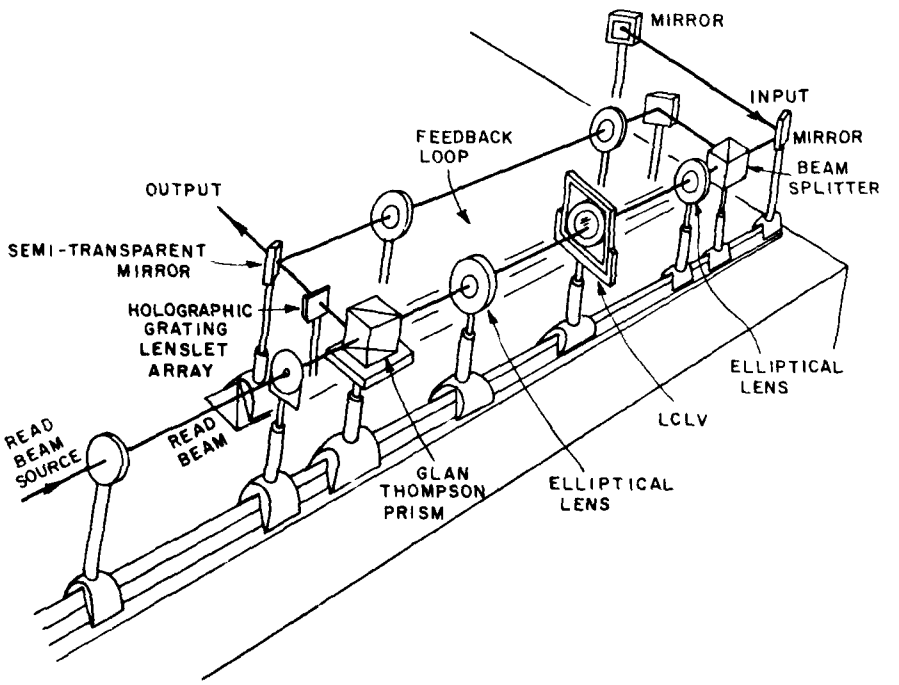


Figure 9. Sketch showing storage apparatus for the temporal integrator.

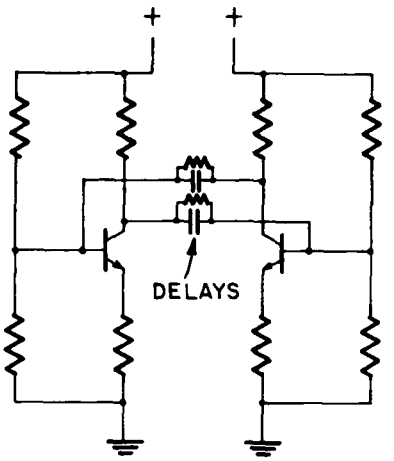


Figure 10. Equivalent electrical circuit.

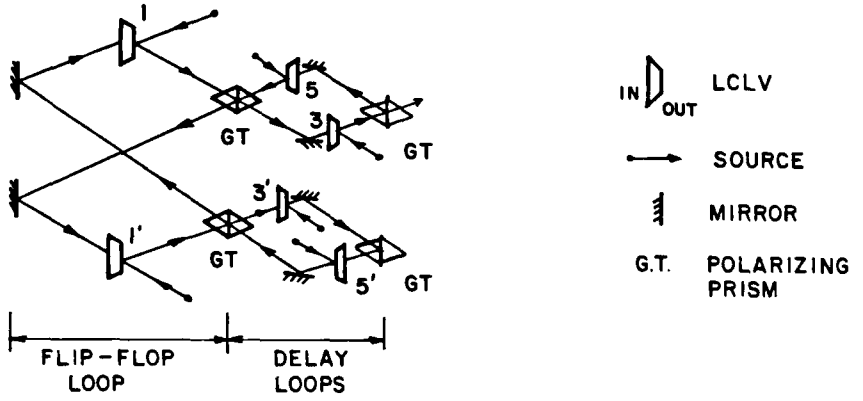


Figure 11a. Basic configuration.

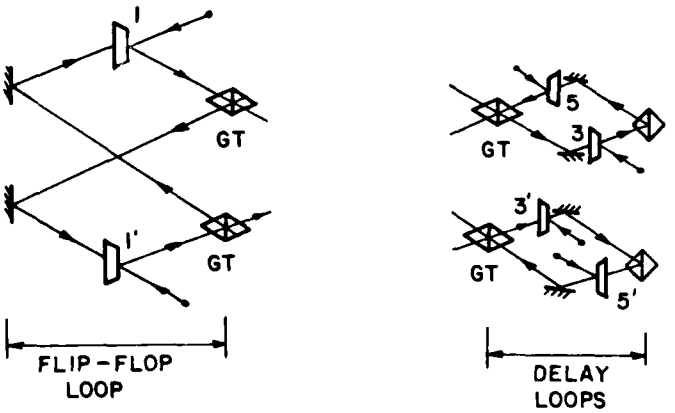


Figure 11b. Sketch of flip-flop loop.
11c. Sketch of delay loops.

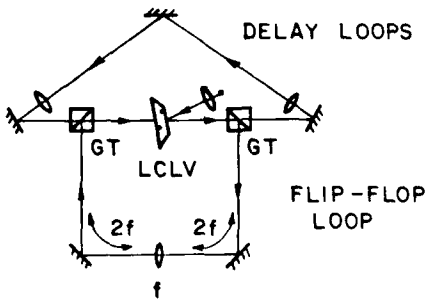


Figure 12. Schematic illustration of optical configuration.

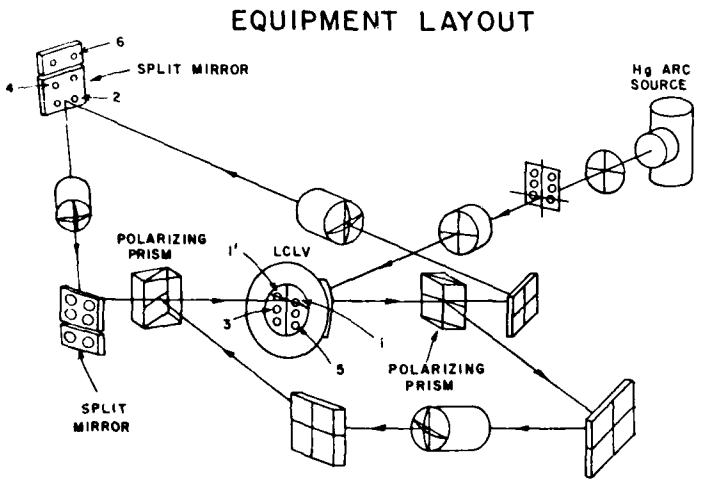


Figure 13. Sketch showing optical layout.

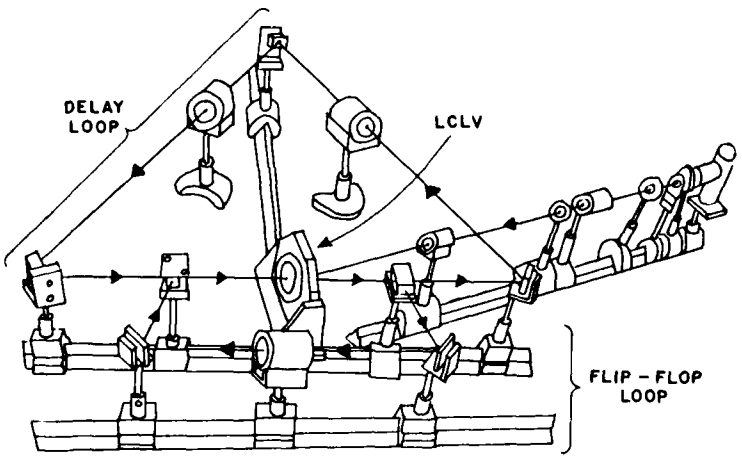


Figure 14. Sketch of actual apparatus.

FORREST, J. Dr‡	Dept. of Electronic and Electrical Engineering, University College London, Torrington Place, London WC1E 7JE, United Kingdom
GARDINER, P.E. Mr	R. & P.E. Dept. Smiths Industries, Bishops Cleeve, Cheltenham, GL52 4SF, United Kingdom
GERHARDT, L. Prof.*	ECSE Dept., R.P.I., Troy, N.Y. 12180, United States
GHICOPOULOS, B. Dr*	H.A.E., Technology Research Centre (KETA), Delta Falirou, P.Faliron, Athens, Greece
GIBBS, H. Dr‡	Optical Sciences Center, University of Arizona, Tucson, AZ 85721, United States
GLASGOW, J.A. Mr	Marconi Avionics (ESL), Basildon, Essex, United Kingdom
GOLZIO, D. Mr	Aeritalia-Gruppo Sistemi Avionici ed Equipaggiamenti, 100072 Caselle (TO), Italy
GRUNEWALD, K. Dr	INT, Appelsgarten 2, 5350 Euskirchen, Federal Republic of Germany
HALL, E.L. Mr	Marconi Avionics Ltd., Airport Works, Rochester, Kent, United Kingdom
HARRISON, R.G. Dr‡	Dept. of Physics, Heriot-Watt University, Riccarton, Edinburgh EH14 4AS, United Kingdom
HAUS, J.W. Dr‡	8814 Louis Dr., Huntsville, AL 35802, United States
HOFFMANN, K. Mr	Bundesamt fur Wehrtechnik und Beschaffung, BWB ML, Dachauerstrasse 128, 8000 Munchen 19, Federal Republic of Germany
HOHLBEIN, K. Mr	Messerschmitt-Bolkow-Blohm GmbH, Postfach 801160, AE-43333, D-8000 Munchen 80, Federal Republic of Germany
HOIVIK, L. Dr*	N.D.R.E., P.O. Box 25, N-2007 Kjeller, Norway
HUNT, G.H. Dr* (Panel Deputy Chairman)	R.A.E., Farnborough, Hants, GU14 6TD, United Kingdom
INCE, A.N. Prof. Dr*	Science and Engineering Council, Burumcuk Sok 7-10, Ankara, Turkey
JACOBSEN, M. Mr*	AEG Telefunken, A14V3, Postfach 1730, D-7900 Ulm, Federal Republic of Germany
KASCHUFTZ, H. Mr	Bundesamt fur Wehrtechnik und Beschaffung, BWB ML, Dachauerstrasse 128, 8000 Munchen 19, Federal Republic of Germany
KASDIUTZ, H. Dr	Messerschmitt-Bolkow-Blohm GmbH, Postfach 901160 AE-322, D-800 Munchen 80, Federal Republic of Germany
KERSTEN, R. TEL Prof.	Fraunhofer Gesellschaft, IPM, Heidenhofstrasse 8, 7800 Freiburg, Federal Republic of Germany
KEYDEL, W. Dr*	DFVLR, NE-HF, D-8031 Oberpfaffenhofen, Federal Republic of Germany
KLEMM, R. Dr*	FEM-FGAN, D-5037 Wachtberg, Federal Republic of Germany
KLINGSHIRN, C. Prof. Dr‡	Universität Frankfurt, Physikaliches Institut, Robert Mayer Str. 2—4, 6000 Frankfurt am Main, Federal Republic of Germany
KUNY, W. Mr*	Messerschmitt-Bolkow-Blohm GmbH, Postfach 801160, LKE4, D-8000 München 80, Federal Republic of Germany
KUNZ, J. Mr	Luftwaffenamt, Postfach 902 500-501 14, 500 Koln, Federal Republic of Germany
LAUG, M. Mr	CERT, B.P. 4025, 31055 Toulouse Cedex, France
LAYCOCK, L.C. Dr	G.E.C. Research Labs, Hirst Research Centre, East Lane, Wembley, HA9 7PP, United Kingdom
LOHMANN, A.W. Prof. Dr‡	Physikalisches Institut, Erwin-Rommel Str. 1, D-8520 Erlangen, Federal Republic of Germany
MACFADYEN, D.N. Dr	Barr & Stroud Ltd, Caxton Street, Anniesland, Glasgow, G12 1HZ, United Kingdom
MACKINTOSH, I.W. Mr*	R.S.R.E., St Andrews Road, Malvern, Worcs, United Kingdom

* Member of Avionics Panel
‡ Author of paper presented at the meeting

LIST OF PARTICIPANTS

ALTMAN, W. Mr	Plessey Telecommunications, Taplow Court, Maidenhead, Berks, UK
BAKKER, L. Ir.	Physics Laboratory TNO, P.O. Box 96864, 25009 JG The Hague, Netherlands
BENNION, I. Mr†	Plessey Research (Caswell) Ltd., Allan Clark Research Centre, Caswell, Towcester, Northants NN12 8EQ, United Kingdom
BERNABE, B.A. Dr	Dept of Army, Avionics R & D Activity, SAVAA-T, Ft Monmouth, N.J. 07760, United States
BIED-CHARRETON, P. Mr	Thomson CSF, 178 Bld Gabriel Péri, 92240 Malakoff, France
BOGENBERGER, R. Mr	Messerschmitt-Bölkow-Blohm GmbH, Postfach 801160 LKE 326, D-8000 München 80, Federal Republic of Germany
BOLLE, H. Dr	Fraunhofer Institut, IITB, Sebastian-Kneipp-Str 12–14, 7500 Karlsruhe, Federal Republic of Germany
BOSMAN, D. Prof. Ir.*	Twente University of Technology, Dept. of Electrical Engineering, P.O. Box 217, 2509 AE Enschede, Netherlands
BOURBIN, Y.P. Dr†	Thomson CSF, Laboratoire Central de Recherches, Domaine de Corbeville, B.P. 10, 91401 Orsay Cedex, France
BOWDEN, C. Dr†	Research Directorate, US Army Missile Laboratory, Attn: DRSMI-RDD, Redstone Arsenal, AL 35898, United States
BRUNNER, H. Dr	ESG Elektronic-System-GmbH, Postfach 80 05 69, 8000 München 80, Federal Republic of Germany
CARROLL, J.E. Prof.†	Engineering Department, University of Cambridge, Trumpington Street, Cambridge CB2 1PZ, United Kingdom
CHADEAU, A. Mr*	STTE PNI, 129 rue de la Convention, 75731 Paris Cedex 15, France
COLLINS, S.A. Prof.†	Ohio State University, Electroscience Laboratory, 1320 Kinnear Road, Columbus, OH 43212, United States
CORBISIER, F. Lt. Col.*	Etat Major Force Aérienne, Quartier Reine Elisabeth, Rue d'Evere, 1140 Bruxelles, Belgium
CROVELLA, L. Mr*	Aeritalia-Gruppo Sistemi Avionici ed Equipagiamenti, 100072 Caselle (TO), Italy
DAINO, B. Prof.	Fond. Bordoni, I.S.P.T., Viale Europa 160, 00144 Roma, Italy
DAKIN, J. Dr	Plessey Electronic Systems Research, Roke Manor, Romsey, Hants, United Kingdom
DE BERNARDI, C. Dr	CSELT, Via Reiss Romoli 274, 10147 Torino, Italy
DEMARIA, C. Mr	Aeritalia GVC Armament Group, Csa Marche 41, 10146 Torino, Italy
DIAMOND, F.L. Dr* (Panel Chairman)	Chief Scientist, RADC/CA Rome Air Development Center, AFSC Griffiss AFB, N.Y. 13441, United States
DOREY, J. Mr*	ONERA, 29, Ave de la Division Leclerc, 92320 Châtillon sous Bagneux, France
DOVE, B. Mr*	MS 469, NASA Langley Research Ct., Hampton, VA 23665, United States
DREXEL, U. Mr	Bundesamt für Wehrtechnik und Beschaffung, BWB ML, Dachauerstrasse 128, 8000 München 19, Federal Republic of Germany
EKRE, H. Mr*	N.D.R.E., P.O. Box 25, N-2007 Kjeller, Norway
FEVROT, C. Mr	49, rue de la Bienfaisance, 94300 Vincennes, France
FISHER, A.D. Dr	Code 6537, Naval Research Lab., Washington, D.C. 20375, United States

* Member of Avionics Panel
† Author of paper presented at the meeting

A.Lohmann, Ge

I think one way for advancing the art of optical computing is to find problems where the electronic people are in trouble and only we can solve the problems. So, for example, one project we have worked on involved triple correlations. Triple correlations are very good for some signal processing purposes, but they are very computationally intensive. They can be done optically, so that is one example. So encourage everyone to come up with problems that are difficult for electronics that helps us to get more support!

The Hopfield model, which Hyatt Gibbs just mentioned, is one of those things. The Hopfield model is a model of the human memory. It even includes a feature of learning and the feature of slowly forgetting and relearning and other things. It's basically a parallel model. The parallelism of that particular model is essential. It consists of a looping system — a data mixing and interconnection step and then a nonlinear step which does simply hard clipping or almost hard clipping. So it's a project which Demetri Psaltis realized can be very easily implemented optically or at least it has good chances so I think that's one of the things we should work on. People are always talking about these super-knowledge base systems of the future for which MITI wants to spend so much money. Well, the Hopfield model is a good indicator of a problem one should work on: optics has a good chance of doing some of these interesting futuristic projects.

H.Gibbs, US

I may be in danger of having my intelligence considered to be artificial in this subject, but it seems to me that A.I. is an important application. We're beginning to go in that direction. The evolution of processors may not necessarily be towards digital computation, but it may be towards image recognition. Hopfield has a theory of how the brain actually operates with massive interconnections, and Psaltis is trying to construct a model using optics that would mimic this kind of thing. Somehow man decides whether it's a friend or an enemy that comes into the cave rather quickly and not in a serial kind of fashion. I think to the extent that optical computers lend themselves to this kind of massive parallel comparisons and conclusions it may be that that will be its eventual real strength.

D.Bosman, Ne

Maybe I didn't make myself clear. When I say that a hundred samples per second is sufficient, that means complete calculations. If one uses a pipeline structure where a number of processors do subsequent processing, and if this pipeline structure is 12 to 15 processors long, then if you would care to minimize size and weight, etc., then you might go for a higher data rate, multiplexing all the separate functions in one general purpose type of computer. So internally in the computer it might be 20 times faster than what your actual input and output rate is.

Unidentified Speaker

And the requirements and accuracy inside the computer and inside the algorithm, depending very much on what you are going to do, might be much higher than the accuracy of your input data.

S.D.Smith, UK

But the encouraging thing about the comments is that Hyatt and I would say the interference filter devices will be fast enough to do things 100 times per second. I would say we could process a complete picture through several stages at 100 times a second. That means — it's very important — that means that we don't have to go for nanosecond time scales in the first stage — we can work with microsecond time scales which looks like a much more practical proposition. So that's a very, very valuable comment from the pair of you, I think.

A.Lohmann, Ge

I have another comment that's been mentioned by, I think, Professor Lohmann and yourself as well about the famous von Neumann bottleneck which is usually referred to. The amount of parallel operations, parallel cables and whatsoever could be reduced by this. I try to defend von Neumann. There's another point: the way the von Neumann computer is thinking is exactly the way human beings are used to thinking, and maybe there's another point that inhibits the development of all these parallel structures: namely, that the human brain is not capable of exploiting these parallel structures for general purposes.

Unidentified Speaker

I'd like to comment on the comparisons which are being made, or have been made this afternoon between optical processing and electronic processing. These have been almost entirely on the basis of whether we need the additional power — or speed — of an optical processor. It seems to me that for many avionic applications there is another characteristic which is equally, or more important, and that is the integrity of computation. I think this is fairly obvious — there are many applications where integrity is of enormous importance. So really the question I'm asking is what will the optical computer offer us in terms of improved integrity, either because the basic elements of which the computer is manufactured are themselves inherently of very high integrity or because there is such a lot of parallelism you can build in enormous safeguards against malfunction. Either of those two, I think, would offer some really interesting potential applications.

J.P.Dakin, UK

Professor Lohmann satisfied my previous question about the diffraction but he raised another question relative to integrity and that is: these Fabry-Perot devices are by their nature highly reflective and when you cascade numerous systems with lenses to focus these in a forward direction clearly reflected energy will be focused back in the reverse direction. There will be numerous of these cascaded with numerous opportunities for levels to build up. Has anyone analyzed the effects of reflection in such systems, especially cascaded systems?

A.Lohmann, Ge

Yes, there are some isolators that can solve this problem; one is a Faraday trap which you use in connection with polarization. A Faraday trap is a device where light goes through in one direction but cannot return the other direction. So inserting Faraday light traps would solve this problem, or at least contribute to the solution of this problem.

S.D.Smith, UK

There's another way too you can go about it — Andy Walker talked about off-axis address and working in reflection and one can avoid the beam coming back — in fact, I was going to just quote: somebody said we ought to have an opponent in electronics. We do have. Robert W.Keys writes excellent articles at regular intervals saying it's all rubbish, but he makes remarks like that — that optical components will be liable to beams coming back and interfering with things. Well you can certainly arrange the geometry so that that doesn't happen, although it isn't trivial, and of course there are Faraday Isolators. He also made remarks, for example in his paper, he said transistors all so easily accommodate varying fan-out and fan-in and offer excellent isolation of input from output. Optical logic elements do not share these desirable qualities. I suggest that that's absolutely wrong. So I mean our opponent is being extremely useful in putting up statements like that which were true in concept a year ago, but aren't true anymore. We can see our way round these completely and indeed fan-out and fan-in may be brilliant in optics.

Unidentified Speaker

I would like to comment on this. If you are talking on, say, intelligent methods for adaptive methods for general or clutter or whatever supression these methods are very sensitive and the problems increase with increasing order of the matrix. You know the problem — you are talking about 1000 × 1000 matrix and try to invert this I think it's a hard job, either with electronic or optical computers, so be careful. On the other hand, there are concepts for radars with omnidirectional transmission and these require parallel forming of all the beams and this is also very good application for optical systems.

Unidentified Speaker

Generally in numerical computation the faster the computer computes the more complicated a computation somebody wants to do — (they want to do a bigger matrix —); the bigger the computation you do, generally the larger the wordlength you need because there's more opportunity to be subtracting numbers that are of similar size and getting into roundoff errors, so the faster you make the machine the longer the wordlength requirements. So people who do modelling on Vax computers almost invariably do everything in double precision. A 32-bit floating point is not all they want; they need the extra size. To some extent, as you make the computer faster, you are going to need longer wordlengths. People would like to have the flexibility to change the algorithms even in future weapon systems so there is some interest in having the floating point even in signal processing application even though there may be a penalty, a severe penalty, in performance at the time, it then gives flexibility to change all those algorithms for more advanced algorithms next year without redoing a lot of equipment.

S.D.Smith, UK

May I say just one more word about this — Professor Lohmann raised the question of inputting. I believe we've already reached the point whereby showing that you can actually get incoherent-coherent conversion in these nonlinear devices which is likely to expand and spread amongst them, that you really can put an awful lot of information into such an array which we can be sure can be done now. You can also put it in with a diode array if you want to, and you can also put it in with some sort of optical scan so that inputting looks, in an initial set of data, very possible. I'd like to re-raise Hyatt's point. Having put it in once, what sort of cycling rate interests people to do things with it; and is there any military interest in putting something like an image into something like this? Then what would you like to do with it?

Unidentified Speaker

I'm speaking for the radar community. We would be happy to have complex matrix algebra of the well-known algorithms at a rate of about 1 to several megahertz.

S.D.Smith, UK

So you'd like complex matrix algebra at a rate of about a megahertz. That's an interesting comment already and quite helpful.

H.Gibbs, US

I'm not sure how that translates — I mean, if we talk about one of these interference filters a million pixels in parallel going at 10 kHz then you can do perhaps 10^{10} bits per second but I don't know whether that's useful or not, because I'm not sure what you're going to do with it...

J.Forrest, UK

Can I come back on that one too — in the sense that I think the systems people need to do a lot of thinking. It's not only fundamental thought on the optical logic because I do worry about the situation where we could get hold of this kind of processor and we end up by digitizing everything in sight with a vast number of parallel beams. And if we compare with what the human eye does, it actually has only one high resolution agile beam. It has a very wide field of view with relatively low resolution processing which may be enhanced in certain ways. It uses this information very intelligently so as not to overload subsequent processors. I think we need from the system angle not just to think: here is something that could give us 10^{11} bits per second. I think the key, as you mentioned Mr Chairman, is in the parallelism and perhaps in using this intelligently. We may not indeed want the 16 bits — we may be quite satisfied with 1 or 2 bits — but at a fairly rapid rate and perhaps with some degree of preprocessing done in the system to ease the burden on subsequent parts of the chain.

Unidentified Speaker

I think that a lot of computation can be done at the very slow rate. When you present images to the pilot, it is not necessary to compute more than let's say 50 times per second. For missile guidance: I think that slow rates may be sufficient around 100 times per second or a little more, but it may be not necessary to think about very high rates.

D.Bosman, Ne

I would like to amplify the previous speaker's statement. Not talking about a radar environment where you might want to supervise more than 100 aircraft simultaneously, if you look at the aircraft as such it has its own vehicle dynamics which will determine the sampling rate of the calculations. I agree that about 100 samples per second — but not just pixels, complete pictures. It's sufficient.

S.D.Smith, UK

Perhaps I can add to Professor Lohmann's remarks by saying that if you do it with holograms it is worth noting that using so-called degenerate-four-wave mixing you have a real time hologram and so that could be programmed in principle. There are various ways in which that could be done. That function itself could be optically programmed and that's certainly feasible at the moment. That's perhaps worth considering.

H.Gibbs, US

In terms of trying to develop a parallel optical processor, I'd like to ask the question of how many pixels in parallel do you need to be interesting? Is it 10^4, 10^5, or 10^6 and in addition what cycle time, what rate do you really need? We talked about a 100 MHz rate, but maybe a kHz is interesting? Are there people involved enough in parallel signal processing that they could guide those of use who come from optical logic individual devices in what we should really be thinking about?

S.D.Smith, UK

Well, Hyatt's point was similar to that one which we wrote down. I just put it up to emphasize it. Can anyone from the military or electronics side or data processing side say what they think would be an interesting demonstration? I think his point is a very good one.

E.Spitz, Fr

The data rate, for instance, in a phased-array radar is usually, say several MHz per channel, and the number of channels can vary between several hundred and several thousand. That's a typical data input rate.

A.W.Lohmann, Ge

There are basically two types of parallelism. You are referring to one-dimensional parallelism which is very suitable if you are dealing with temporal data which you for example want to convolve continuously as is the case for radar signal processing. In image processing two-dimensional parallel processing is desirable because semantically the connection between different pixels is of a two-dimensional nature. I would say typically something that can be considered, something like an image or at least which is large enough to include a total image feature, is maybe of the order of 128 × 128 pixels, so if the first optical computer can handle 128 pixels squared then it can really handle something that has recognizable pictorial features. And the time constant in the first instance maybe should not be very short because you run into many synchronization problems which you would probably like to avoid in the first run. On the other hand, of course, nonlinearities get stronger if you compact your energy both in time and space.

H.Gibbs, US

What's the time?

A.W.Lohmann, Ge

Well, a hundred MHz rate is quite all right for the beginning.

Powers

I think in regard to many of the phased array requirements at present one could say that with the sytems that are being conceived or are being built there isn't quite the requirement for this processing rate. One is talking about relatively few multiple beams and not vast numbers of elements. But, of course, one of the big requirements for the future is the possibility that one could rapidly adapt beam patterns particularly to cancel out interference. And the traditional way of adapting such an antenna pattern would be to change the phase and amplitude distributions right across the whole aperture. So one is then asking, effectively in real time, and one is talking about the time constants at which interference sources can change which is well under the microsecond level, to work out a new distribution for perhaps thousands of elements. The classical algorithms on this involve matrix inversion, so unless one can find better ways of doing this you are talking about inverting vast matrices in real time, matrices which are thousand by thousand elements. The kind of processing rates that this requires are beyond the current digital capabilities and are certainly into this sort of range. So here it does strike me as a real requirement that might use this kind of technology. There has been consideration of course, by a number of groups — Dave Cassesant has talked about this sort of application, for analog optical processors. But the problem there, of course you can indeed do matrix inversions of this kind of size very rapidly, but not accurately enough and this strikes me as the point where you really do need the enhanced dynamic range or accuracy that you could get from a *digital* optical system.

S.D.Smith, UK

Before you go away, could you say a little bit more about this about what this would imply from what you've heard of what we are doing — this question of the high dynamic range? How many digits do we need and that sort of thing?

Powers

Yes. Well I think you're really talking about 16 bits for many of these systems. That's just to give a ballpark figure here, because to get the kind of sidelobe levels for antennas that are typically discussed, like 40 dB or better, one is talking about precisions of the order of a few degrees of phase, or sometimes better and a few tenths of a dB in amplitude. So you are talking about 1% or better and this leads you into these sorts of orders of magnitude of 16 bit processing.

EEC PROJECT DISCUSSION

Unidentified Speaker

To me x^1 is something which has real meaning only for processes in which the time you are interested in is long compared to any natural resonance time in the medium. It is extremely relevant if your are talking about soliton propagation in fibres where we know the effects on the picosecond time domain are still slow compared to the response of the fibre. You can model everything, and it works out beautifully. But it does seem to me that in many of the applications in which you are using it, that is true for the experiments that are done at almost D.C.; but it is not necessarily true at the rates for which you would wish to use the device, since the actual processes that you are invoking to produce what I would say is a pseudo x^1 effect are, in fact, perfectly real transitions with real changes in absorption spectra and in refractive index. It might be better to simply stick to that as a notation which then automatically highlights the temporal dependence of those effects, leading directly into a device modelling which will tell whether or not it will work in a particular regime.

S.D.Smith, UK

Should I respond? Yes, he is quite right. The only real reason for introducing the x^1 is to make comparison, I think. And of course there is a lot to be said about the time scales. It gets interesting when the time scale gives the possibility of a coherent response of the system and one would have to have a proper description in terms of x^1, but for device purposes you can lump it together as a nonlinear index. You can usually compare by just turning it into the dimensions of x^1. But your comment is quite right. The time scales have to be fed in properly and defined.

Unidentified Speaker

Can I ask everyone a question on the forefront of making real devices from nonlinear etalons? I raised the question yesterday in my talk about whether or not we want to avoid threshold logic. An awful lot of people seem to talk about making AND gates by just adding signals together to the level that they will switch. Talking to electronics people, this is just not the way you should do it. Can one rule that approach out totally, or would there be circumstances in which threshold logic would be an acceptable approach to designing a useful machine? Has anyone got any response to that?

Unidentified Speaker

Let me offer one trivial response. It surely depends on how accurately you have control of your amplitudes, i.e., if the signals have already been hard limited before you go through the add and threshold stage, it's fine — if you haven't, it isn't.

Unidentified Speaker

Yes, that obviously is the point, and of course in electronics, one avoids having to be that fussy about levels very often. One point that might be made is, in terms of the levels of signals coming into gates, there is a distinct difference in an optical system in fan-out compared to an electrical one in which the signal is defined by a parallel. If you choose to fan-out to four devices, you've got to divide your power by four. If you choose to fan-out to only two, then you've got to have twice as much. At some stage you've got to get these all back together to some standard levels; certainly if you are sticking to threshold logic that causes problems. May I just conclude that point by saying that I think we can get round that and use non-threshold approaches.

J.P.Dakin, UK

I'm new to the field, from the outside looking in. People have used terms like fan-out and basic mechanisms of switching beams. However, we have to remember here that this first stage produces an extremely narrow beam which if there is a high density of bistable states is an extremely narrow beam originally of parallel form. To try and produce holographic methods of switching this involves operating over an extremely small area and there are problems in getting the necessary resolution with any form of basic grating effect to actually deflect the beam in the required direction. It it's a very narrow beam, it will produce a very broad deflected beam. The pipeline structure may work, but there will clearly be diffraction effects after the first stage of processing — from all the various optical levels involved over a very small array. And I think the basic diffraction optics are something which people haven't really considered here. I think perhaps I would like to hear more discussion on it.

A.W.Lohmann, Ge

May I respond to this and possibly go to the blackboard? It is easier to make a painting for a pipeline processor, although the feedback processor isn't really different in principle. One way to look at it: you might have here a data plane... and you have a lens that forms from every point an image to the next data plane. So this is data plane number one, and this is data plane number two, and this is data plane number three, and so on. And at every plane, here is some logic going on — an array of logic components. Now a hologram or some other component that does the redirection of light beams, for example, these beams are now deflected this way so that they are ending up at this point instead. So if this is hologram no.1 or deflector no.1 etc. then you would have here the next one and for this you would need another lens which opticians would call a field lens. This field lens, then, would image this point by means of these rays into one point which is then here again. So again there is an imaging from deflector planes to deflector planes so you have a nested system of imaging — one side here and the other side here — this nested system of imaging is so that you conserve the number of pixels throughout the total pipeline system (of course instead of doing pipelines one could roll it up into a feedback system — this would be basically the same) but in this sense you really can maintain the number of pixels, let's say hopefully one million or so, over several periods.

DISCUSSION OF CLOSELY COUPLED TWIN STRIPE LASERS WITH BISTABILITY

J.W.Haus, US

Could you make any comment about other active bistable devices like these cleave-coupled cavity devices or others that have been made?

J.E.Carroll, UK

They have their cavities in parallel; we have ours in tandem. And whether in fact they're modal bistability or whether they're using more dispersive bistability this idea was — certainly devices in tandem — was first thought of by Basov a long time ago and developed or constructed by Lasher. It didn't work too well. They relied on absorptive bistability trying to get the one level so there wasn't sufficient light for lasing and another level so there's sufficient light in both elements to maintain lasing. The cleaved-coupled cavity variety I think is of that sort rather than the modal bistability which I think we're operating here. In the cleave-couple cavity I believe we're going to have a change in electron concentration, and I suspect that will therefore be slower. But time will tell.

DISCUSSION

S.D.Smith, UK

That was a super finish! Could you say just exactly what you ask the liquid crystal light valve to do in your operations? Is it amplifying? Is it stepping?

Author's Reply

Basically it is a controlled birefringent mirror with roughly 360,000 elements; a polarizer and analyzer convert it to a polarization rotator or an amplitude changing device whichever I want.

S.D.Smith, UK

Well, does that mean it was acting as a gate?

Author's Reply

No. To make it act as a gate what I'll do is bring in a single beam reflected off the light valve twice with a combination of polarizers and analyzers in there and find that in order to have it get through the whole set of polarizers and analyzers I have to have specific inputs. Otherwise I get zero out and that's what a gate does.

S.D.Smith, UK

Sure, I understood the gate. What was it doing when it wasn't a gate?

Author's Reply

OK. When it wasn't a gate it was an optical amplifier in the case of the residue adder; in the case of the master slave flip-flop it was a flip-flop.

S.D.Smith, UK

It was sometimes an amplifier and sometimes a flip-flop.

Author's Reply

Basically it's your active element, your transistor, and as you make various transistor circuits, so you make various light valve circuits. Except that there's the equivalent of 360,000 transistors.

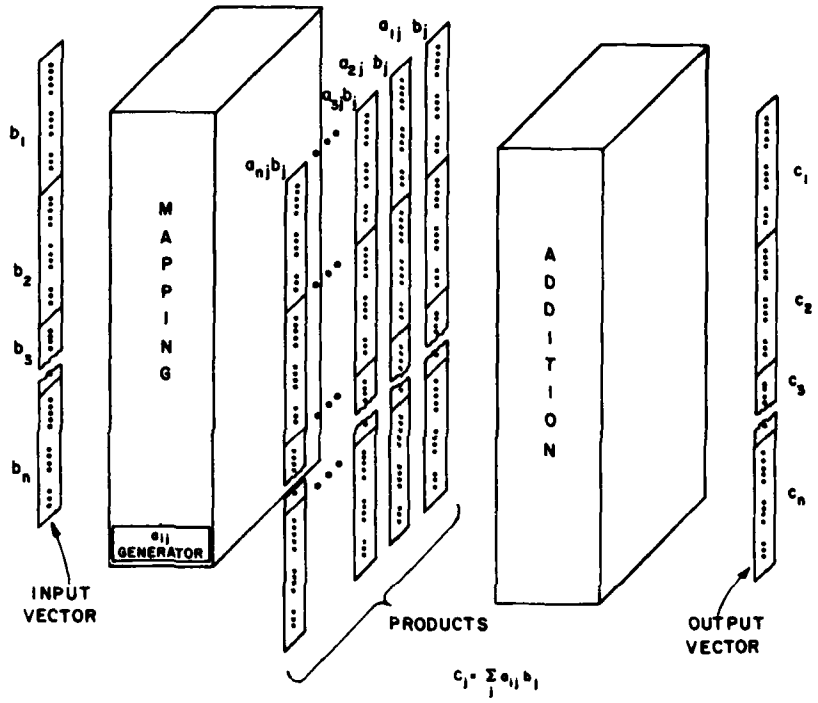


Figure 23. Block diagram of matrix vector multiplier.

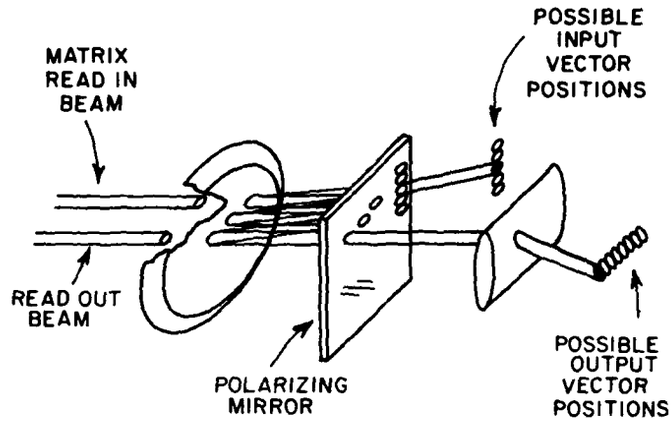


Figure 21. Optical apparatus for multiplication mapping implementation.

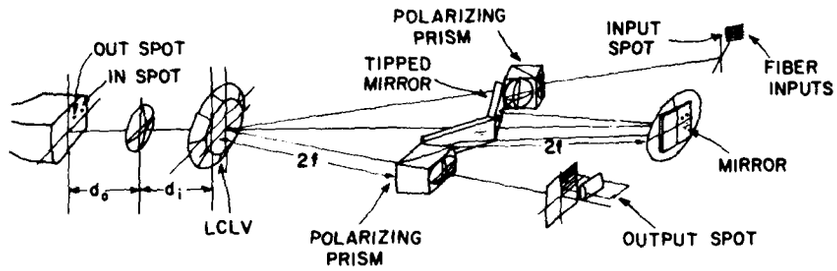


Figure 22. Detail of mapping implementation in multiplication unit.

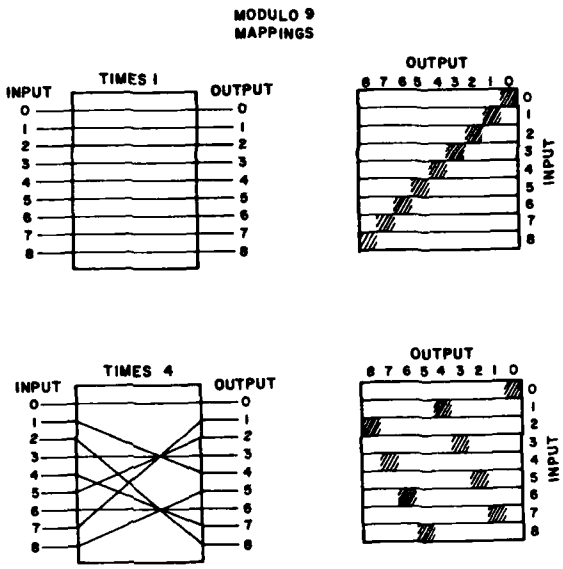


Figure 19. Modulo 9 truth table mappings for multiplication by one and four.

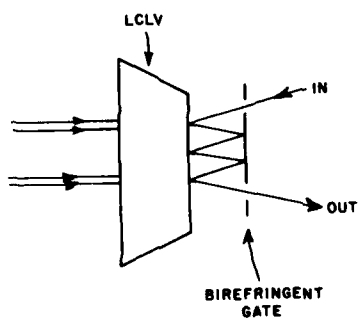


Figure 20. Birefringent gate: mapping implementation principle.

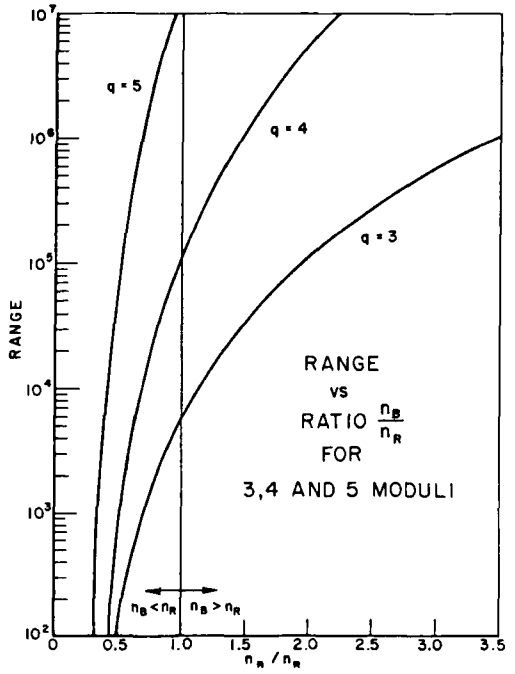


Figure 17. Plots of range as a function of the ratio of binary to residue temporal integrators, for 3,4, and 5 base representations.

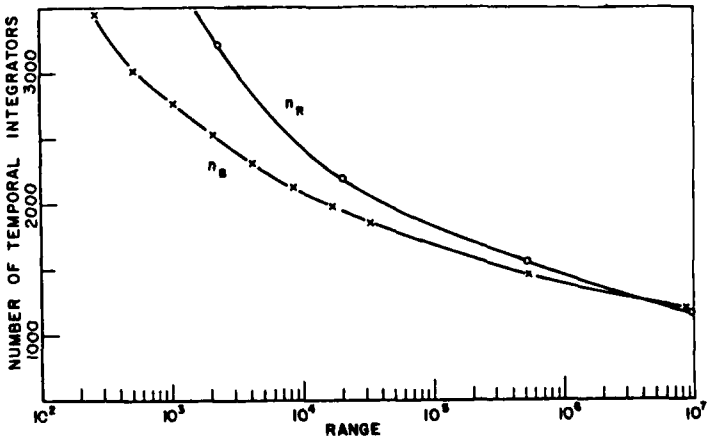


Figure 18. Plot of the number of binary and residue temporal integrators as a function of range.

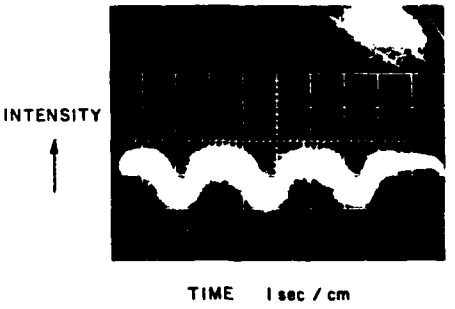


Figure 15. Photograph of oscilloscope trace of clock output.

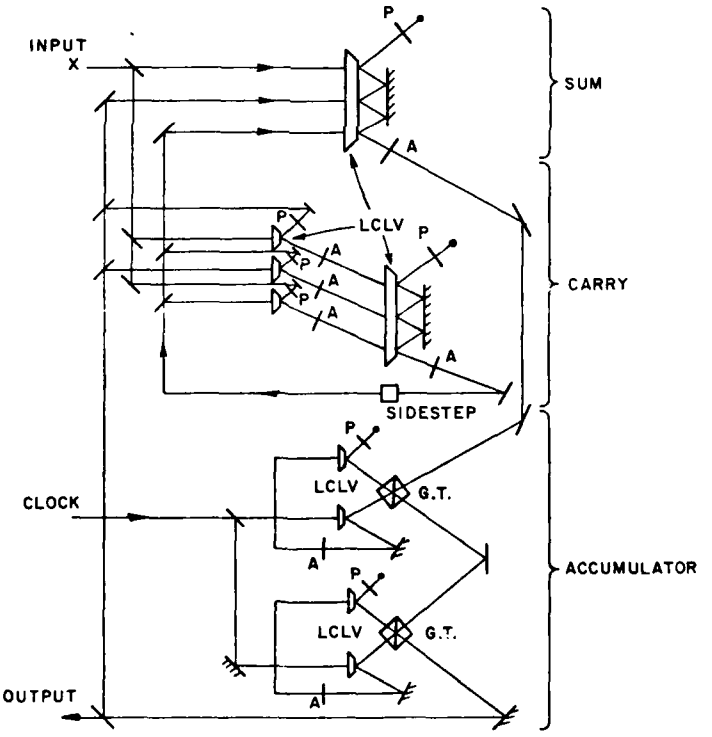


Figure 16. Optical logic-gate-based binary temporal integrator.

MACPHERSON, R.W. Dr*	National Defence Headquarters, CRAD DSP-3, 101 Colonel By Drive, Ottawa, Ontario K1A 0K2, Canada
MALARD, M. Mr	Thomson CSF, 52, rue Guynemer, 92130 Issy-les-Moulineaux, France
MASCARENHAS, J.M.B.G. Maj. Eng.*	Head, Test Software Group, O.G.M.A., 2615 Alverca, Portugal
MCAULAY, A.D. Dr†	Texas Instruments Inc., Central Research Laboratories, P.O. Box 266015, MS 238, Dallas, TX 75266, United States
MIDWINTER, J.E. Prof.	Dept. of Elec. Eng., University College, Torrington Place, London W1E 7AG, United Kingdom
MILLER, A. Dr†	R.S.R.E. Malvern, Worcs, WR14 3PS, United Kingdom
MONTALTI, F. Dr	Industrie Face Standard Research Centre, Via Della Magione 10, 00040 Pomezia, Italy
NEFF, J. Dr†	DARPA/DSO, 1400 Wilson Blvd., Arlington, VA 22209, United States
OYENUGA, O. Mr	Pb 256, British Aerospace, P.O. Box 19, Six Hills Way, Stevenage, SG1 2DA, United Kingdom
PATTONI, S. Mr	Ferranti plc, Dunsinane Ave, Dundee DD2 3PN, United Kingdom
REICHART, G. Mr	Luftwaffenamt, Postfach 902 500/501 14, 5000 Köln, Federal Republic of Germany
REISCHMANN, H.L. Mr	Bodenseewerk Geraetetechnik, FFK-IR, Postfach 11 20, 7770 Überlingen, Federal Republic of Germany
REMONDINI, A. T. Col.	Com. do Corpo Tecnico Esercito, Via A. Scarpa 18, 00161 Roma, Italy
RITCHIE, S. Mr†	British Telecom Research Labs., R4.4.3, Martlesham Heath, Ipswich, IP5 7RE, Suffolk, United Kingdom
RYLES, J.C. Dr*	Chief Scientist, AFWAL/AS, Wright Patterson AFB, OH 45433, United States
SALAUN, M. Mr	Avions Marcel Dassault, 78 Quai Carnot, 92 St Cloud, France
SCHEPERS, H.J. Dr	DFVLR e.V., ZPK, Postfach 90 60 58, 5000 Köln 90, Federal Republic of Germany
SCHIEBSCHICK, L. Mr*	Deputy Director STC, SHAPE Technical Centre, P.O. Box 174, 2501 CD The Hague, Netherlands
SCHIEK, R. Mr	Institut für Technische Elektrophysik, Arcistrasse 21, 8 München 2, Federal Republic of Germany
SCHLIEN, G. Mr	Bundesamt für Wehrtechnik und Beschaffung BWB ML, Dachauerstrasse 128, 8000 München 19, Federal Republic of Germany
SMITH, S.D. Prof.†	Dept. of Physics, Heriot-Watt University, Riccarton, Edinburgh EH14 4AS, United Kingdom
SPITZ, E. Dr	Directeur Technique et de la Recherche, Thomson, 173 Bld Haussmann, 75008 Paris, France
STROLE, D. Dr	Siemens AG, Postfach 3240 D-852 Erlangen, Federal Republic of Germany
SUETA, T.J. Mr*	Dept. of Army, Avionics R & D Activity, Fort Monmouth, N.J. 07760, United States
TAILLEBOIS, J. Mr	MCB, 11, rue Pierre Lhomme, 92404 Courbevoie, France
TALLET, A. Dr	Batiment 213, Univ. Paris-Sud, 91405 Orsay Cedex, France
TEGTMEIER, H.D. Mr	Messerschmitt-Bölkow-Blohm GmbH, Postfach 801160, LKE 4, D-8000 München 80, Federal Republic of Germany
TISCHER, S. Dr	6803 Edingen, Anna-Sender 7, Federal Republic of Germany
VAN BOCHOVE, A.C. Ir.	Dr Neher Laboratories, St Paulusstraat 4, 2264 XZ Leidschendam, Netherlands
VAN DEN BRANDEN, W. Mr*	Etat Major Force Aérienne, Quartier Reine Elisabeth, Rue d'Evère, 1140 Bruxelles, Belgium
VEITH, G. Dr†	SEL Research Center (ZT/FZNG), H.Hirth Str. 42, D-7000 Stuttgart 40, Federal Republic of Germany

* Member of Avionics Panel
† Author of paper presented at the meeting

VERPY, J.J. Mr	Sintra, 26, Rue Malakoff, 92600 Asnières, France
WALKER, A.C. Dr†	Dept. Physics, Heriot-Watt University, Riccarton, Edinburgh, United Kingdom
WALKER, N.G. Mr†	Cambridge University, Engineering Dept., Trumpington St., Cambridge, United Kingdom
WARD, J.J. Mr	Smiths Industries, Bishops Cleeve, Gloucestershire, United Kingdom
WEIERHOLT, A.J. Mr	ELAB-NTH, 2034 Trondheim, Norway
WEISS, M.T. Dr*	The Aerospace Corp. MS:M1-002, P.O. Box 92957, Los Angeles, CA 9000, United States
WHALLEY, J. Mr*	British Aerospace Plc., Weybridge Div. Woodford, Bramhall, Stockport, Cheshire SK7 1QR, United Kingdom
WINTERS, B.H. Dr	Laboratory for Physical Sciences, 4928 College Ave. College Pk, Maryland 20740, United States

* Member of Avionics Panel
† Author of paper presented at the meeting

REPORT DOCUMENTATION PAGE

1. Recipient's Reference	2. Originator's Reference	3. Further Reference	4. Security Classification of Document
	AGARD-CP-362	ISBN 92-835-1485-8	UNCLASSIFIED

5. Originator Advisory Group for Aerospace Research and Development
North Atlantic Treaty Organization
7 rue Ancelle, 92200 Neuilly sur Seine, France

6. Title
DIGITAL OPTICAL CIRCUIT TECHNOLOGY

7. Presented at
the 48th Meeting/Specialists' Meeting of the Avionics Panel held in Schliersee, Federal Republic of Germany, 11—12 September 1984.

8. Author(s)/Editor(s)
Billy L.Dove (Editor)

9. Date
March 1985

10. Author's/Editor's Address
NASA Langley Research Center
Mail Stop 469
Hampton Va 23665, USA

11. Pages
234

12. Distribution Statement
This document is distributed in accordance with AGARD policies and regulations, which are outlined on the Outside Back Covers of all AGARD publications.

13. Keywords/Descriptors

Digital data processing	Optical computer	Optical modulators
Digital optics	Optical demodulators	Optical sources
Electro-optics	Optical detectors	Optical waveguides
Integrated optics	Optical devices	Parallel processing
Optical bistability	Optical logic	

14. Abstract

These Proceedings for the 48th Meeting of the AGARD Avionics Panel contain the 18 papers presented, a Technical Evaluation Report, and discussions that followed the presentations of papers. Seven papers were presented in the session devoted to optical bistability. Optical logic was addressed by three papers. The session on sources, modulators and demodulators presented three papers. Five papers were given in the final session on all optical systems.

The purpose of this Specialists' Meeting was to present the research and development status of digital optical circuit technology and to examine its relevance in the broad context of digital processing, communication, radar, avionics and flight control systems implementation.

AGARD Conference Proceedings No.362 Advisory Group for Aerospace Research and Development, NATO DIGITAL OPTICAL CIRCUIT TECHNOLOGY by Billy L.Dove Published March 1985 234 pages These Proceedings for the 48th Meeting of the AGARD Avionics Panel contain the 18 papers presented, a Technical Evaluation Report, and discussions that followed the presentations of papers. Seven papers were presented in the session devoted to optical bistability. Optical logic was addressed by three papers. The session on sources, modulators and demodulators presented three papers. Five papers were given in the final session on all P.T.O	AGARD-CP-362 Digital data processing Digital optics Electro-optics Integrated optics Optical bistability Optical computer Optical demodulators Optical detectors Optical devices Optical logic Optical modulators Optical sources Optical waveguides Parallel processing	AGARD Conference Proceedings No.362 Advisory Group for Aerospace Research and Development, NATO DIGITAL OPTICAL CIRCUIT TECHNOLOGY by Billy L.Dove Published March 1985 234 pages These Proceedings for the 48th Meeting of the AGARD Avionics Panel contain the 18 papers presented, a Technical Evaluation Report, and discussions that followed the presentations of papers. Seven papers were presented in the session devoted to optical bistability. Optical logic was addressed by three papers. The session on sources, modulators and demodulators presented three papers. Five papers were given in the final session on all P.T.O	AGARD-CP-362 Digital data processing Digital optics Electro-optics Integrated optics Optical bistability Optical computer Optical demodulators Optical detectors Optical devices Optical logic Optical modulators Optical sources Optical waveguides Parallel processing
AGARD Conference Proceedings No.362 Advisory Group for Aerospace Research and Development, NATO DIGITAL OPTICAL CIRCUIT TECHNOLOGY by Billy L.Dove Published March 1985 234 pages These Proceedings for the 48th Meeting of the AGARD Avionics Panel contain the 18 papers presented, a Technical Evaluation Report, and discussions that followed the presentations of papers. Seven papers were presented in the session devoted to optical bistability. Optical logic was addressed by three papers. The session on sources, modulators and demodulators presented three papers. Five papers were given in the final session on all P.T.O	AGARD-CP-362 Digital data processing Digital optics Electro-optics Integrated optics Optical bistability Optical computer Optical demodulators Optical detectors Optical devices Optical logic Optical modulators Optical sources Optical waveguides Parallel processing	AGARD Conference Proceedings No.362 Advisory Group for Aerospace Research and Development, NATO DIGITAL OPTICAL CIRCUIT TECHNOLOGY by Billy L.Dove Published March 1985 234 pages These Proceedings for the 48th Meeting of the AGARD Avionics Panel contain the 18 papers presented, a Technical Evaluation Report, and discussions that followed the presentations of papers. Seven papers were presented in the session devoted to optical bistability. Optical logic was addressed by three papers. The session on sources, modulators and demodulators presented three papers. Five papers were given in the final session on all P.T.O	AGARD-CP-362 Digital data processing Digital optics Electro-optics Integrated optics Optical bistability Optical computer Optical demodulators Optical detectors Optical devices Optical logic Optical modulators Optical sources Optical waveguides Parallel processing

optical systems.

The purpose of this Specialists' Meeting was to present the research and development status of digital optical circuit technology and to examine its relevance in the broad context of digital flight control systems implementation.

Papers presented at the Avionics Panel Symposium held in Schliersee, Germany, 11—12 September 1984.

ISBN 92-835-1485-8

---

optical systems.

The purpose of this Specialists' Meeting was to present the research and development status of digital optical circuit technology and to examine its relevance in the broad context of digital flight control systems implementation.

Papers presented at the Avionics Panel Symposium held in Schliersee, Germany, 11—12 September 1984.

ISBN 92-835-1485-8

---

optical systems.

The purpose of this Specialists' Meeting was to present the research and development status of digital optical circuit technology and to examine its relevance in the broad context of digital flight control systems implementation.

Papers presented at the Avionics Panel Symposium held in Schliersee, Germany, 11—12 September 1984.

ISBN 92-835-1485-8

---

optical systems.

The purpose of this Specialists' Meeting was to present the research and development status of digital optical circuit technology and to examine its relevance in the broad context of digital flight control systems implementation.

Papers presented at the Avionics Panel Symposium held in Schliersee, Germany, 11—12 September 1984.

ISBN 92-835-1485-8

END

DATE
FILMED

8 -85

DTIC